```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

May 11, 2004, 11:39:45; Search time 120.24 Seconds (without alignments) 1738.900 Million cell updates/sec Run on:

US-09-917-376-3

4036 1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

1586107 seqs, 282547505 residues Searched:

1586107 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

A_Geneseq_29Jan04:* .. Database

geneseqp2003as:* geneseqp2003bs:* geneseqp1980s:* geneseqp1990s:* geneseqp2002s:* geneseqp2000s:* geneseqp2001s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

geneseqp2004s:*

SUMMARIES

į		æ,			SOFTWARES	
Result No.	Score	Query	Query Match Length	DB	I.D	Description
	4036	100.0	740	9	ABP73016	Abp73016 Amino aci
7	4036	100.0	740	7	ADD22926	Add22926 Acidother
m	4036	100.0	740	7	ADD22923	Add22923 Acidother
4	4036	100.0	957	9	ABP73015	Abp73015 Amino aci
2	4036	100.0	957	7	ADD22921	Add22921 Acidother
9	2063	51.1	940	9	ABB99489	Abb99489 Amino aci
7	1680	41.6	726	7	ADD22927	Add22927 Aspergill
80	1579	39.1	838	7	ADD42055	
σ	841.5	20.8	555	7	ADD24919	Add24919 Xanthomon
10	238	5.9	2468	9	ABU38411	Abu38411 Protein e
11	238	5.9	2468	9	ABP59933	Abp59933 Microbial
12	199.5	4.9	2435	9	ABU19529	Abu19529 Protein e
13	181.5	4.5	3892	9	ADA34216	Ada34216 Acinetoba
14	181	4.5	1308	9	ABU16642	Abu16642 Protein e
15	170	4.2	1074	9	ABU22692	Abu22692 Protein e
16	168.5	4.2	1439	ø	ABU16643	Abu16643 Protein e
17	168.5	4.2	5291	7	ADC01014	Adc01014 Enterohae
18	167.5	4.2	1119	7	ABW01170	Abw01170.Candida a
19	166	4.1	1468	7	ADD48744	Add48744 Rat Prote
20	163.5	4.1	1289	9	ABU16977	Abu16977 Protein e
21	163	4.0	1684	~	AAR14948	Aar14948 Bacterial
22	160.5	4.0	3716	9	ABM15900	Abm15900 Mycobacte
23	159.5	4.0	1291	~	AAW59912	Aaw59912 Amino aci
24	158.5	3.9	774	~	AAW35390	Aaw35390 Flavobact
25	158.5	3.9	1440	Ŋ	ABB54801	Abb54801 Lactococc

Abu36440 Protein e Abp43908 MUC5B par Ade08753 Novel pro Abr65361 Rifidhac		Abulysos Florein e Aae37932 Human CGD Abu34624 Protein e Aay90913 Cenarchae	Aay78516 Isoamylas Abu17029 Protein e Abp69842 Human pol Abn69841 Human pol	
ABU36440 ABP43908 ADE08753 ABP6541	AAY13493 AAE16324 AAP90615	ABU19388 AAE37932 ABU34624 AAY90913	AAY78516 ABU17029 ABP69842 ARP69841	ABP69840 AAU31850 ABU34548 ABU39844 AAP94419
957	122	9 2 9 8	ოდიი	.u4001
2204 4315 1531	1751	827 1488 1721 3472	776 1325 1419	1516 2367 1106 3619 750
 		, n n n n a a a	8. E. E. E	
156.5 154.5 153	152.5 152.5 152.5	152 152 152	151.5 151 150.5	150.5 150.5 149.5 149
26 27 28	22220	2 2 2 2 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5	3 8 8 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

ALIGNMENTS

ABP73016 standard; peptide; 740 AA. (first entry) 03-JUN-2003 ABP73016; RESULT 1 ABP7301

Amino acid sequence of the avicelase AviIII catalytic domain.

Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile; cellulose.

Acidothermus cellulolyticus.

İ

WO2003012090-A2.

13-FEB-2003.

28-JUL-2001; 2001WO-US023818.

28-JUL-2001; 2001WO-US023818.

(MIDE) MIDWEST RES INST.

Himmel ME; Ding S, Adney WS, Vinzant TB,

WPI; 2003-248177/24.

New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.

Claim 6; Page 8; 44pp; English.

The present sequence is derived from a thermostable avicelase, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing not textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polynucleotide is useful as a

```
740; Conservative
                                                                                                                                                   WPI; 2003-810853/76.
                                                                                                                  HIMMEL M E.
                                                                                         DING S.
ADNEY W S.
VINZANT T E
                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                             Sequence 740 AA;
                         US2003108988-A1
                                          12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                         (VINZ/)
(HIMM/)
                                                                                                                                 ŝ
                                                                                         (DING/)
(ADNE/)
                                                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                    Ding
                                                                                                                                                                                                                                                                                                                                             Matches
  g
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                  ö
                                                                                                  120
                                                                                                                 120
                                                                                                                                  180
                                                                                                                                                  180
                                                                                                                                                                   240
                                                                                                                                                                                   240
                                                                                                                                                                                                   300
                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                   360
                                                                                                                                                                                                                                                    360
                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                                      540
                                                                                                                                                                                                                                                                                                                                                     540
                                                                                                                                                                                                                                                                                                                                                                      600
                                                                                                                                                                                                                                                                                                                                                                                      900
                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                      720
                                                                                                                                                                                                                                                                                                                                                                                                                                                      720
                                                                                                                                                                                                                                                                                                                                                                                                                      999
                                                                  9
                                                                                 9
                                                                                                           TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                            SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                              ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                  NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTY1ANPTDT
                                                                                                                                                                                                                                   SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                    APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                     AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                              QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                   TGYQSDIQGVVWVAFDKSSSLGQASKT1FVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                   IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKPSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                    PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                             PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                              APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                      WAPGDPGOPVVYAVGFGNSWAASOGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                     WAPGDPGOPVVYAVGFGNSWAASOGVPANAOIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                      KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                  Gaps
                                                  ö
                                 Length
                                                  Indels
primers in various diagnostic assays
                                Score 4036; DB 6;
Pred. No. 8.4e-271;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               standard; protein; 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YIGTNGRGIVYGDIGGAPSG
                                 100.0%;
                          Query Match
Best Local Similarity 100.0
**** A0; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
  or
 probes
                 Sequence 740 AA;
oĘ
                                                                                                                 61
                                                                                                                                  121
                                                                                                                                                  121
                                                                                                                                                                  181
                                                                                                                                                                                  181
                                                                                                                                                                                                   241
                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                   481
                                                                                                                                                                                                                                                                                                                                                                      541
                                                                                                                                                                                                                                                                                                                                                                                     541
                                                                                                                                                                                                                                                                                                                                                                                                                     601
                                                                                                                                                                                                                                                                                                                                                                                                                                                      661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      721
                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                   421
                                                                                                                                                                                                                                                                                                                                                                                                      601
                                                                                                                                                                                                                                                                                                                                                                                                                                     661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADD22926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADD22926;
  sonrce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADD22926
SXS
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                            셤
                                                                                                8
                                                                                                             셤
                                                                                                                                 ઠ
                                                                                                                                               셤
                                                                                                                                                                 ò
                                                                                                                                                                                셤
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                  dd
                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

180 480 300 300 360 420 420 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG 120 360 New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste. ø NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKKLGG NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIANPTDT SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY 1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW NMPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY 241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI The invention relates to an isolated polynucleotide molecule encoding thermostable Avilli polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding Avilli. The polynucleotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. Method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase Avilli catalytic domain. Gaps ö Length 740; Indels 100.0%; Score 4036; DB 7; 100.0%; Pred. No. 8.4e-271; ive 0; Mismatches 0; Himmel ME Example 2; SEQ ID NO 6; 29pp; English. Adney WS, 'Vinzant TB, 28-JUL-2001; 2001US-00917376. 18-OCT-2002; 2002US-00155400 cellulolyticus 361 421 셤

ö

120

9 9

180 240 240

Acidothermus cellulolyticus avicelase AviIII catalytic domain.

enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.

ö

Gaps

9 9 120 120 180 180 240 240 300 300 360 360 420 420 480 480 540 540 900 600 99 99 720

8

à

```
KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Avicelase, AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile; cellulose.
                                                                                                                                             121 NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDT
                                                                                                                                                                                                                                                                                                                                                                                TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                        IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                        SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                          PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                         481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       601 OPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KSAPGSSYPAVFVVGT1GGVTGAYRSDDCGTTWVL1NDDQHQYGNWGQA1TGDHANLRRV
                                                                                              1 ATTOPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWWGW
                                                                                                                               NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKT,GG
                                                                                                                                                                                                                                                TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   421 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                     ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                       NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                          ö
               Length
                                          Indels
            100.0%; Score 4036; DB 7;
100.0%; Pred. No. 8.4e-271;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Amino acid sequence of the avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABP73015 standard; protein; 957 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                          Conservative
                        al Similarity
740; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  199
             Query Match
Best Local S
Matches 740
                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              721
                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            601
                                                                                                                                 61
                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                         241
                                                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABP73015
                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
                                                                                                                                                                                                                                                                           쉽
                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                               ð
                                                                                                                                                                                       ò
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                         099
                                                         540
                                                                                    900
                                                                                                                900
                                                                                                                                             9
                                                                                                                                                                                                                       KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV 720
APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to an isolated polynucleotide molecule encoding thermostable AvilII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucleotide is useful for reducing cellulose in starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII catalytic domain.
                                                                                                    AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                             QPVAAGLPSSGAVGVMFHAVPCKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                    WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                              AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Acidothermus cellulolyticus avicelase AviIII catalytic domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Himmel ME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 16; SEQ ID NO 3; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                   ADD22923 standard; protein; 740 AA
                                                                                                                                                                                                                                                                              YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                          15-JAN-2004 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DING S.
ADNEY W S.
VINZANT T B.
HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2003-810853/76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 740 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12-JUN-2003
                            481
                                                       181
                                                                                    541
                                                                                                                                                                                                      199
                                                                                                                                                                                                                                   661
                                                                                                                                                                                                                                                               721
                                                                                                                                             601
                                                                                                                                                                                                                                                                                            721
                                                                                                                                                                                                                                                                                                                                                                                               ADD22923;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (DING/)
(ADNE/)
(VINZ/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (HIMM/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ding
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                         g
                                                                                                                                           충
                                                                                                                                                                        g
                                                                                                                                                                                                8
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                       셤
```

480

996

```
New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                        QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to an isolated polynucleotide molecule encoding thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is
                                   467 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                             AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                            WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                    QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                         KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                       AELNPSI IVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       enzyme, AviIII; cellulose reduction, agricultural biomass; municipal solid waste, glycoside hydrolase, avicelase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus avicelase AviIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Œ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Himmel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 16; SEQ ID NO 1; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Æ
                                                                                                                                                                                                                                                                                                                                                                YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                  786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADD22921 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /label= Unknown
                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001US-00917376.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2003-810853/76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DING S.
ADNEY W S.
VINZANT T E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; ADD22922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Key
Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12-JUN-2003
                                                                                                         527
                                                                                                                                                                                587
                                                                                                                                                                                                                                                     647
                                                                                                                                                                                                                                                                                                                             707
                                                                                                                                                                                                                                                                                                                                                                                                  767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADD22921;
 421
                                                                       481
                                                                                                                                              541
                                                                                                                                                                                                                    601
                                                                                                                                                                                                                                                                                         661
                                                                                                                                                                                                                                                                                                                                                                721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (ADNE/)
(VINZ/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (DING/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (HIMM/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADD2292
                                                                                               ď
                                                                                                                                                                            g
                                                                                                                                                                                                                                              쉽
                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                          Š
                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                             ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence represents a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII su useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-aviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   466
                                                                                                                                                                                                                                                                                                                                          New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 4036; DB 6; Length 957; 100.0%; Pred. No. 1.2e-270;
                                   ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                               /note= "unspecified residue encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                     Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                     Himmel
Location/Qualifiers
                                                                                                                                                                                                                                                     TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 2; Page 20; 44pp; English.
                                                                                                                                          28-JUL-2001; 2001WO-US023818
                                                                                                                                                                              28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                     Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 100.
nes 740; Conservative
                                                                                                                                                                                                                (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                       WPI; 2003-248177/24.
                957
                                                                                                                                                                                                                                                     Adney WS,
                                                                                                                                                                                                                                                                                                        N-PSDB; ABZ77632.
               Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 957 AA;
                                                                   WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                     တွဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                     Ding
```

ò g ò ପ୍ର ò g ò 임 ò 셤 ò g ઠે

```
03-OCT-2002
                                                               Jonesia sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             104
                                                                                                                                                                                                                                                                                                                                                                                                                                       φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              305
                                                                                                                                                                              Duffner
                                                                                                                                                                                                                                                                Claim
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                         360
                                                                                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                                                       540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       766
                                                                                                                                                                     120
                                                                                                                                                                                                           180
                                                                                                                                                                                                                            226
                                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                                                    241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY 300
                                                                                                                                                                                                                                                                                                      346
                                                                                                                                                                                                                                                                                                                                          406
                                                                                                                                                                                                                                                                                                                                                                                466
                                                                                                                                                                                                                                                                                                                                                                                                                    526
                                                                                                                                                                                                                                                                                                                                                                                                                                                         586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   706
                                                                                                                                                 106
                                                                                                                                                                                        166
                                                                                                                                                                                                                                                                 286
                                                                                                                                9
                                                                                                                                            ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                      NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                               NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                     NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                  APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                       AELNPSI IVRAGSFDPSSQPNDRHVAFSTDGGXNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA1TGVSSAVNVGFG
                                                                                                                                                                                                                                               TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                          I PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                              PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
      administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII.
                                                                                                               Gaps
for reducing cellulose in a starting material which involves
                                                                                                              ö
                                                                                            Length
                                                                                                              Indels
                                                                                          ; Score 4036; DB 7;
; Pred. No. 1.2e-270;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YIGTNGRGIVYGDIGGAPSG
                                                                                           Query Match
Best Local Similarity 100.0%;
Matches 740; Conservative 0
                                                                          Sequence 957 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   647
                                                                                                                                                                     61
                                                                                                                                                                                                           121
                                                                                                                                                                                                                             167
                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                  227
                                                                                                                                                                                                                                                                                                      287
                                                                                                                                                                                                                                                                                                                          301
                                                                                                                                                                                                                                                                                                                                           347
                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                407
                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                     467
                                                                                                                                                                                                                                                                                                                                                                                                                                       481
                                                                                                                                                                                                                                                                                                                                                                                                                                                          527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
ABB99489
ID ABB9
XX
AC ABB9
8$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                            Б
                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                8 8
                                                                                                                                                                    Š
                                                                                                                                                                                  유
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                               ò
```

```
125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGILSLATDPVDTNRVYLAAGTYSMDWDPQNGAILRSADKGETWEKTMLPFRVGGNMPGR 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMV 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IDRQHPNTIMVATQISWMPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLT 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence represents a xyloglucanase enzyme, belonging to family 74 of glycosyl hydrolases. The enzyme is isolated from Jonesia sp. DSM14140. The enzyme is useful in processes for machine treatment of fabrics. It is also useful in the textile industry for improving the properties of cellulosic fibers, yarn, woven or non-woven fabric, and in the cellulose fiber process step. The xyloglucanase enzyme is also useful in the cellulose fiber processing industry for ratting of fibers such as hemp, jute, flax and linen. It is useful for preventing binding of certain soils to the xyloglucan left on the cellulosic material
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTG-YQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLT 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New xyloglucanase enzyme belonging to family 74 of glycosyl hydrolases, and endogenous to a bacterium, useful in the textile industry for improving properties of cellulosic fibers, yarn or fabric.
                                                                                                                     fiber;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    164 GMGERLAVDPNNNKVLYFGAESGNGLWKSTDYGKTWGKVTSFPNAGNYVA---DASGAYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       221 GQNQGVVWVTFDPTSAKAGQTTQT1YVGVADKQNNVYRSTDGGATWQRVPGQPTGFLAQK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YTWSNVAIGGGGFVDGIVPNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18; Gaps
                                                                                                                     cellulosic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 940;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                               Xyloglucanase; enzyme; family 74; glycosyl hydrolase;
textile scouring; cellulose fiber processing; ratting.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 51.1%; Score 2063; DB 6; Best Local Similarity 52.1%; Pred. No. 5.2e-134; Matches 386; Conservative 112; Mismatches 225;
                                                             of a xyloglucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5; Page 72-75; 76pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                 27-MAR-2001; 2001DK-00000504.
                                                                                                                                                                                                                                                                                                                                                                                         2002WO-DK000210
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F, Sjoholm C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2003-092855/08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (NOVO ) NOVOZYMES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; ABV76941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 940 AA;
                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

400

standard; protein; 940

ABB99489

ABB99489

```
avicelase III catalytic domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
SXS
                                                                                                                                                         g
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                         ò
                                                                                                                                      ð
                                                                                                                                                                                  ð
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
           483
                                 519
                                                        543
                                                                            577
                                                                                                   603
                                                                                                                           635
                                                                                                                                                  657
                                                                                                                                                                       693
                                                                                                                                                                                              717
                                                                                                                                                                                                                     753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to an isolated polynucleotide molecule encoding thermostable Avili polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding Avilii. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Aspergillus aculeatus
                     --EGAAPRRSTNSGSSWSTVSGLPHNAQVASDRVNANTLYGFVDGKFYHSTNGGASFTAS
                                                                                                                                                                                                          544 GDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV
          KGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPST-IFTSPVFTTGTSVDYAEL
                                                     184 NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAP
                                                                            520 KPATVVRAGK--SISGETTSWVGVSTDAGETWKPGATPSGVKGPGSITVSANASSIVWAP
                                                                                                                                                 A-AGLPSSGAVGVMFHAVPGKEGDLWLA----ASSGLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                              GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANL
                                                                                                                                                                                                                                                                                                                                                                                                                    agricultural biomass;
                                                                                                                                                                                                                                                                                                                                                                                                                   enzyme; AviIII; cellulose reduction; agricultural bion
municipal solid waste; glycoside hydrolase; avicelase.
                                                                                                                                                                                                                                                                                                                                                                                            Aspergillus aculeatus avicelase III catalytic domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MΕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Himme1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 2; SEQ ID NO 7; 29pp; English.
                                                                                                                                                                                                                                           718 RRVYIGTNGRGIVYGDIGGAP 738
                                                                                                                                                                                                                                                        GRVYIGTNGRGIIVGDSSTPP 774
                                                                                                                                                                                                                                                                                                                         ADD22927 standard; protein; 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Aspergillus aculeatus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADNEY W S.
VINZANT T B.
HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-810853/76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JS2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DING S.
                                                                                                                                                                                                                                                                                                                                                                     15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-JUN-2003
                                160
                                                                                                                          578
                                                                                                                                                 604
                                                                                                                                                                                             658
                                                                                                                                                                                                                    594
                                                                                                                                                                                                                                                               754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             s,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (DING/)
(ADNE/)
(VINZ/)
(HIMM/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ding
                                                                                                                                                                                                                                                                                                           RESULT 7
                                                                                                                                                                    유
                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                               a
                           유
                                                   ò
                                                                       유
                                                                                                 ò
                                                                                                                      g
                                                                                                                                         ò
                                                                                                                                                                                            ò
                                                                                                                                                                                                                 임
         8
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                            116
                                                                                                                                                                                                                                                                                                                                                                                                                                                   116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DGDTVLLMSSTSGALVSKSQG---TLTAVSSLPSGAVIASDKSDNTVFYGGSAGAIYVSK 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         177 SSST--YTSDPVGIAWVTFDSTSGSSGSATPRIFVGVADAGKSVFKSEDAGATWAWVSGE 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DYFGYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : ||||: | | : | | : | | : | | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Trichoderma reesei; Family 74 xyloglucanase; beta-1,4-glycosidic linkage; xyloglucan hydrolysis; biomass degradation; cellulose; hemicellulose;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT 711
                                                                                                                                                                                                                                                                                              59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PQYGFLPHKGVLSPEEKTLYISYANGAGPYDGTNGTVHKYNITSGVWTDISP---TSLAS
                                                                                                                                                                                                                                                                      PT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ||| ||| : || : || : ||| : ||| ||| ||| DISNAPWIQDITISTDQFPV---RVGMMYEALAIDPFDSNHWLYGTGLIVYGGHDLINWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 ATTOPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                                                                                                                                                                                                                                                                            ---wnnwgyngvvsiaadpintnkvwaavgmytnswdpndgailrssdogatwoitPlpF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISARPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS
                                                                                                                                                 Gaps
                                                                                                                                             32;
                                                                      Length 726;
                                                                                                                                                 Indels
                                                              tch 41.6%; Score 1680; DB 7; al Similarity 46.2%; Pred. No. 1.2e-107; 346; Conservative 113; Mismatches 258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Trichoderma reesei Family 74 xyloglucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDLQTYGRVFRGHERPGHLLRQSQREPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADD42055 standard; protein;
Sequence 726 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADD42055
                                                                      Query Match
                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADD42055
```

The

```
471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               from Xanthomonas campestris which may be used for activity reduction or enhancement using directed genetic engineering. A transformed cell or organism having reduced or enhanced activity of at least one such protein e.g. galactomannanase can be generated by directing the gene encoding the protein. The activity of the protein is reduced by the presence of an antisense nucleic acid sequence. The nucleic acid sequence of the gene
                                                                                                                                                548
                                                                                                                                                                                                                                                                                648
                                                                                                                                                                                                                                                                                                      606 VSKDTGSSF---TRG-PKLGSAGTIRDIAAHPTTAGTLYVSTDVGIFRSTDSGTTFGQVS 661
                                                                                                                                                                                                                                                                                                                                                               706
                                                                                                                                                                                                                                                                                                                                                                                         531
                                                                                                                                                                                            591
                                                                                                                                                                                                                                    605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention relates to polypeptide and polynucleotide sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     transformed cell or organism having reduced or enhanced activity of least one protein, useful for producing xanthan gum, which are useful providing formulations and properties.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Directed genetic engineering; galactomannanase; reduced activity; enhanced activity; xanthan gum production; suspension stability; emulsion stability; temperature resistance; pseudoplasticity; amylase; cellulase; extracellular protease; intracellular protease;
                                                                                                                               415 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHA---DVTAVPSTIFTSPV
                                                             434 RHNVSIQSLADGIEEFSVQDLASAPGGSELLAAVGDDNGFTFASRNDLGTSPQTVWATPT
                                                                                                                                                                                                                  549 YSADGDILWSTASSG---VQRSQFQSFASVSSLPAGAVIASDKKTNSVFYAGSGSTFY
                                                                                                                                                                                                                                                                                                                                                               TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA--YRSDDCGTTWVLINDDQHQYGNW
                                                                                                       FITGISVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                            532 ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY
                                                                                                                                                                                                                                                                             RSTDGGVTFQPVAAGLPSSGAVGVM--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                  707 GQAITGDHANLRRVYIGTNGRGIVY--GDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         718 STKVAGSGSTAGQVÝVGTŇGRGVFÝAQGTVGGGTGG 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; SEQ ID NO 47; 135pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Xanthomonas campestris cellulase #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADD24919 standard; protein; 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             glucose dehydrogenase; enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-AUG-2001; 2001US-00927827.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-MAR-2001; 2001US-0279493P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ramseier TM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Xanthomonas campestris.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BOWER S G.
RAMSEIER T M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2003-625389/59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; ADD24893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US2003036176-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bower SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New tit.
                                                                                                                                                494
                                                                                                                                                                                                                                                                                592
                                                                                                                                                                                                                                                                                                                                                               649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADD24919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BOWE/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RAMS/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADD24919
ID ADD2
                                                                                                                                        g
                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                        New polypeptides having Family 74 xyloglucanase activity, and encoding nucleic acid molecules, useful for degrading cellulose-and hemicellulose containing biomass to ethanol or as a detergent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a Family 74 xyloglucanase (ADD42055) from the fungus Trichoderma reesei (Hypocrea jecorina), and nucleic acids encoding it (e.g., ADD42054). The enzyme hydrolyses the beta-1,4-glycosidic linkages in the backbone of xyloglucan to xyloglucan oligosaccharides. The invention also relates polypeptide sequences at least 70% identical to the enzyme, expression vectors and host cells comprising a nucleic acid of the invention, the recombinant production of the enzyme, and mutant enzymes and the nucleic acids encoding them. The xyloglucanase of the invention can be used in the degradation of cellulose- and hemicellulose-containing biomass to produce ethanol. It can also be used in a detergent composition for treating fabric during a machine washing cycle. The present sequence represents the Trichoderma reesei Family 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGÅTWSQMTNFPDVGTYIANPTDT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADP-NNPVFWSRDGGATWQAVPGAPTG 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NGYNSDKOGLMWYTFDSTSSTTGGATSRIFVGTADNITASVYVSTNAGSTWSAVPGQPGK 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFG 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                300 YSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISA 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 YTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWV----GW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDOGATWOITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 360 EPWLTFG-VOPNPPVPS----PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44; Gaps
production; detergent composition; fabric treatment;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 838;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 39.1%; Score 1579; DB 7; Best Local Similarity 43.0%; Pred. No. 1.5e-100; Matches 325; Conservative 123; Mismatches 264;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 8; SEQ ID NO 2; 96pp; English.
                                                                                                                                                                                                                                                                           (NOVO ) NOVOZYMES BIOTECH INC
                                                                                                                                                                                        17-APR-2003; 2003WO-US011831.
                                                                                                                                                                                                                                 19-APR-2002; 2002US-0373987P.
                                                                                                                                                                                                                                                                                                                                                            WPI; 2003-845528/78.
N-PSDB; ADD42054, ADD42060.
                   treatment; enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                          New polypeptides having
nucleic acid molecules,
                                                                                                                                                                                                                                                                                                                   Zaretsky E,
                                                             Hypocrea jecorina.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 838 AA;
                                                                                                    402003089598-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 xyloglucanase.
                                                                                                                                                30-OCT-2003
                                                                                                                                                                                                                                                                                                                     Michael R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     197
  ethanol
                     cextile
```

셤

6 6 6 6

ò

셤

6 6 6

8

셤

8

```
encoding the protein is a recombinant sequence having at least one mutation as compared to the wild-type gene encoding the protein. The transgenic cell or microrganism are useful for producing xanthan gum, which are useful for providing formulations and properties, such as longterm suspension and emulsion stability in alkaline, acid, and salt solutions, temperature resistance, and pseudoplasticity. The present sequence represents an enzyme relating to the present invention.
                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                      191
                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                          309
                                                                                                                                                                                                                                                                                                                                                                                                                                                        300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AP-----MVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVF 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    519
                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Antisense; prokaryotic essential gene; cell proliferation; drug design
                                                                                                                                                                                                                                                                                                                                                                                                            APQRPLQWWFQDRGLEETVPLDLLSFWAGAHLLSALGDIDGFRHDDLDRV-QLQYAGPRL
                                                                                                                                                                                                                               | :: | | | : | | | | : | | | | DDWNLMGIDAFAVDPADADALYLAAGTYMHE-RAGTAAVLRSFNRGRTFERADLPFKLGG
                                                                                                                                                                                                                                                                                                                                  NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                                                                                                                                   310 RPSHMAGGSDGH-WYLSYGDQPGPDLMAGGALWKFTPAQGRWREISPIPQ-PASGDGFGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PW-TAHATPH-------WMG-ALAIDPFDGNHALFVTGYGIWASRNL------ODFA
                                                                                                                                                                                                              1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                        NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                           TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                   Gaps
                                                                                                                                                                                    41;
                                                                                                                                                     Length 555;
                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Prokaryotic essential gene #23938
                                                                                                                                                  / Match 20.8%; Score 841.5; DB 7;
Local Similarity 37.1%; Pred. No. 1e-49;
hes 186; Conservative 81; Mismatches 193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABU38411 standard; protein; 2468 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTGTSVDYAELNPSIIVRAGS 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | | |:| | | :||:|:
TNGESIDAAGQAPQWVVRSGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAR-2002; 2002WO-US009107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protein encoded by
                                                                                                                          Sequence 555 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200277183-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3-0CT-2002
                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                              192
                                                                                                                                                                                                                                                                                                                                                                                                                       251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABU38411;
                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
   88888888
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
the invention relates to an isolated mucletic actu comprising any one of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation, or that has an activity against a biological pathway or required for proliferation, or that inhibits proliferation or the biological pathway in which a proliferation, or that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profilling a compound's activity; (11) a culture comprising strains in which the gene or which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the culture or which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the collection of an organism. The antisense nucleic acids required for cellular proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this contraction, the partial genes. Note: The sequence data for this contraction, the partial genes. Note: The sequence data for this contraction.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1215 NGTVIVATATDPTGNTGPQAATTVDAVAPPAPVIDPSNGTTISGTAEAGAKVILTDGNGN 1274
                                                                                                                                                                                                                                                                               screening
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .---TG-----FIPHKGVFDPVN 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1188 -DAGGNPIGQVTAD-----GTPLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1084 VNLSNGSSLSG----TAEPGSTVILTD-----GNGNPIAEVTADGSGNWTYTPSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --NG-VVSIAADPINTNKVWAAVGMYTNS-----WDPNDGAILR-SSDQGATWQITPLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FKLGGNMPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N------PIDTTGYQS----DIQGVVWVAFDKSSSSL--GQASKTIFVGVADPN-N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 VAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY----
                                                                                                                                      Zyskind JW;
Xu HH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88; Mismatches 314; Indels 292; Gaps
                                                                                                                                                                                                                                                                        New antisense nucleic acids, useful for identifying proteins or e
for homologous nucleic acids required for cellular proliferation
isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to an isolated nucleic acid comprising any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2468;
                                                                                                                                      Ohlsen KL,
Forsyth RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.9%; Score 238; DB 6; 22.8%; Pred. No. 4.8e-07;
                                                                                                                                    Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO 66335; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ftp.wipo.int/pub/published pct sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          219 PV-FWSRDGGATWQAVPGAP------
                                                                                                                                      Malone C,
Carr GJ,
25-OCT-2001; 2001US-0342923P. 08-FEB-2002; 2002US-00072851. 06-MAR-2002; 2002US-0362699P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 22.8
Matches 205; Conservative
                                                                                         (ELIT-) ELITRA PHARM INC.
                                                                                                                                      Zamudio C,
Trawick JD,
                                                                                                                                                                                                      2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2468 AA;
                                                                                                                                                                                                                             N-PSDB; ACA42281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176
                                                                                                                                      Wang L,
Wall D,
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
1452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TE-ISGTAEAGATVILT-----DGGGN----PIGQATADGSGNWTFTPSTPLANGTVIN 1548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----LSNGTVVNAVAQDAAGNTSGPVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVI 1759
PIGETTADGSGNWSFIPGTPLANGTVVNAVAQDPAGNTGPQGSTTVDAVAPNTPVVNPSN 1334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----VTADGSGNWSFTPGTPLANGSVINALAQDAAGNNSSPTSATVDSLAPAAPV 1657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LTDGNGNPIGQTLADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTSGPASTTVDTVAPAT 1819
                                                                                                                                                                                                           GNLLNGTAEPGSTVTLTDGNGNPIGQTTADGSGNWSFTPGSQLPNGTVVNVTASDAAGNT 1394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :|| : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                     301
                                                                                                                                                                                                                                                                                                                 GL----TIDRQHPNTIMV----ATQISWWPD---TIIFRSTDGG-----ATWTRIWD 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  393 DRMLYGT----GATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISP----PSGAPLI 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----TGGTVAASADGSRFVWAPGDPGQPV--VYAVGFGNSWAASQGVPANAQIRSDRVNP 579
                                                                                                     HVLYIATSNTGGPY----DGSSGDVWKFSVT-SGTW--TRISPVPS----TDTANDYFGYS
                                                                                                                                                                                                                                                                                                                                                                                                                       SLPATITVDSSLPSIPQVDPSNGSVISGTADAGNTIII---TDGNGNPIGQVTADGSGNWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            343 WT-----SYPURSLRYVLDISAEPWL--TFGVQPNPPVPSPKLGWMDEAMAIDPFNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SALGDLGGFTH-----ADVTAVPSTIF--TSPVFTTGTSVDYAELNPSIIVRAGSFDPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KTFYALSNGTFYR-----STDGGV--TFQPVAAGLP---SSGAVGVMFHAVPGKEGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LWLAASSGLYHS-TNGGSSWSAITGV----SSAVNVGFGKSAPGSSYPAVFVVGTIGGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------OGSEPGGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .658 IDPSNGSVIAGTAEAGATVILTDGNGNPIGOVTADGSGN-WSFTPGTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SQPNDRHVAFSTDGGKNWF --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         549
                                                                                                           252
                                                                                                                                                                                                                                                                                                                      302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                     ઠે
```

Microbial resistance gene PA1874 protein. standard; protein; 2468 resistance 18-SEP-2002; 2002WO-US029565 18-SEP-2001; 2001US-0323241P (first entry) (DART-) DARTMOUTH COLLEGE WPI; 2003-468567/44. N-PSDB; ACC59398. Mah T; Biofilm; microbial 402003041483-A2 28-AUG-2003 O'toole GA, 22-MAY-2003 ABP59933 ABP59933 ABP59933

Identifying modulators of microbial resistance of organisms in biofilms, e.g. inhibitor of biofilm formation, by employing expression controls, or efflux pumps containing polypeptides, of genes associated with biofilm

Claim 33; Fig 7; 102pp; English.

The present invention relates to a method of identifying a compound capable of altering the sensitivity of a microorganism to an antimicrobial agent by employing efflux pumps compurising polypeptides encoded by the following genes: PA1874, PA4142, PA2389, PA1876, PA4143, PA2390 or PA163. The method is useful for identifying modulators of microbial resistance of an organism in a biofilm. The methods are also useful for identifying genes that encode proteins that play a role in biofilm resistance. The method is particularly useful for screening compounds or discovering compositions that will inhibit biofilm formation and overcome their resistance mechanisms. These methods are particularly useful in medical, industrial or natural settings, where formation of biofilms can have serious negative consequences and result in high costs both in human health and economic terms. The present sequence is a protein shown in the exemplification of the invention

Sequence 2468 AA

6 6 6 6 6 6 6

1500 TE-ISGTAEAGATVILT-----DGGGN----PIGQATADGSGNWTFTPSTPLANGTVIN 1548 1395 SLPATTTVDSSLPSIPQVDPSNGSVİSGTADAGNİİI--İDGNGNPIGQVTADGSGNWS 1452 50; 1131 IANGTVVNVVAQDASGNSSPPATVTVDSSAPPAPVINFSNGVVISGTAEAGATVTLT--- 1187 GNILNGTAEPGSTVTLIDGNGNPIGQTTADGSGNWSFIPGSQLPNGTVVNVTASDAAGNT 1394 1605 GQ-----VTADGSGNWSFTPGTPLANGSVINALAQDAAGNNSSPTSATVDSLAPAAPV 1657 252 HVLYIATSNTGGPY---DGSSGDVWKFSVT-SGTW--TRISPVPS----TDTANDYFGYS 301 302 GL----TIDRQHPNTIMV----ATQISWWPD---TIIFRSTDGG------ATWTRIWD 342 WT-----SYPNRSLRYVLDISAEPWL--TFGVQPNPPVPSPKLGWMDEAMAIDPFNS 392 N------PIDTIGYOS----DIQGVVWVAFDKSSSL--GQASKTIFVGVADPN-N -----FIPHKGVFDPVN 1084 VNLSNGSSLSG----TAEPGSTVILTD-----GNGNPIAEVTADGSGNWTYTPSTP --NG-VVSIAADPINTNKVWAAVGMYTNS-----WDPNDGAILR-SSDQGATWQITPLP 116 FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIA 1275 PIGETTADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTGPQGSTTVDAVAPNTPVVNPSN DRMLYGT---GATLYATNDLTKWDSGQIHIAPMVKGLEETAVNDLISP----PSGAPLI SALGDLGGFTH-----ADVTAVPSTIF--TSPVFTTGTSVDYAELNPSIIVRAGSFDPS VAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY----/ Match 5.9%; Score 218; DB 6; Length 2468; Local Similarity 22.8%; Pred. No. 4.8e-07; Local Sonservative 88; Mismatches 314; Indels 292; Gaps 1188 -DAGGNPIGQVTAD-----GSGNWSFTP----PV-FWSRDGGATWQAVPGAP-SQPNDRHVAFSTDGGKNWF-. 1549 219 343 176 393 446 Query Match 99

g

8 8 8

g

8 8 8

유

g

ò

----TGGTVAASADGSRFVWAPGDPGQPV--VYAVGFGNSWAASQGVPANAQIRSDRVNP 579

526

셤

ò

```
the foll antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

(I) a vector comprising a promoter operably linked to the nucleic acid encoding a promoter operably linked to the nucleic acid comprising a promoter operably linked to the nucleic acid conciding a promoter operably linked to the nucleic acid conciding a propertied by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated or polypeptide or its fragment whose expression is inhibited by the antisense antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation, or that has an acclivity against a biological pathway capaired for proliferation, or that inhibits cellular proliferation; (8) dentifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a
                                                                                                                                                                                                                                                                                   1819
                                                                                                 626
                                                                                                                                                        ----LSNGTVVNAVAQDAAGNTSGPVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVI 1759
                                                                                                                                                                                                                         681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 antisense nucleic acids, useful for identifying proteins or screening homologous nucleic acids required for cellular proliferation to late candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                                  1859
                                                                                                                                                                                                                                                                                                                                                    740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Antisense; prokaryotic essential gene; cell proliferation; drug design.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zyskind JW;
Xu HH;
                                                                                              KTFYALSNGTFYR-----STDGGV--TFQPVAAGLP---SSGAVGVMFHAVPGKEGD
                                                                                                                                                                                                                      LWLAASSGLYHS-TNGGSSWSAITGV----SSAVNVGFGKSAPGSSYPAVFVVGTIGGVT
                                                                                                                                                                                                                                                                                LTDGNGNPIGQTLADGSGNWSFTPGTPLANGTVVNAVAODPAGNTSGPASTTVDTVAPAT
                                                                                                                                                                                                                                                                                                                                                                                                               ----PVINPSN-----GSVITGT-AEVGAKVILTDGNGNPIGETTADGSG
                                                                                                                                                                                                                                                                                                                                                 GAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             relates to an isolated nucleic acid comprising any one
ŖŖ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ohlsen K
Forsyth F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Protein encoded by Prokaryotic essential gene #5056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 25; SEQ ID NO 47453; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABU19529 standard; protein; 2435 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malone C,
Carr GJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2001US-00815242.
2001US-00948993.
2001US-0342923P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-MAR-2002; 2002US-0362699P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-MAR-2002; 2002WO-US009107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2002US-00072851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (ELIT-) ELITRA PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2003-029926/
N-PSDB; ACA23399.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Borrelia cepacia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200277183-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-MAR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            36-SEP-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-OCT-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38-FEB-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33-OCT-2002
                                                                                              580
                                                                                                                                                        1705
                                                                                                                                                                                                                      627
                                                                                                                                                                                                                                                                                   1760
                                                                                                                                                                                                                                                                                                                                                    682
                                                                                                                                                                                                                                                                                                                                                                                                           1820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABU19529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , a d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  isolate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABU19529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM OF THE STANFORM O
                                                                                          ð
                                                                                                                                           g
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                          ठे
                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
42;
                                                                          proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of pattent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTN---FP 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPG 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292 TVTATDAAGNVSTPSAAFELTIDTTAPALPTVNATDGTSLSGTAEAGATVNIDTNGDGTP 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DVGTYIANPTDTTGYQSD----IQGVVWV-AFDKSSSSLGQASKTIFVGVADPNNPVFWS 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D-ATVTADPSGVWTYTPSTPLPIGTVIGVTATDAAGNTGPSASVTVTGDTTAPGAPVI-- 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --GTVTDDA--GSVVGAIASGGSTDDATPTL-SGTAEAGSTVSVYDGTTLLGTTTADPSG 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DVWKFSVTSG--TWTRISPVPSTDTANDYFGYS---GLTIDRQHPNTIMV-ATQISWWPD 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------GTVIGVTAT-----DAAGNTGPSASV-----TVTGDTTAP--GAPV 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          445 ISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAG-----SFDPS 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOPNDRHVAFSTDGGKNWF-----QGSEPGGVT---TGGTVAASADGSRFVWAPGDP 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    G-----NWTFTPTTALGEGAHSLTVTATDTAGNVSVPSTAFDLTIDTTAP 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GQPVVYAV-GFGNSWAASQCVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAA 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    697 AIPTVNATDGTSLSGTAEAGATVNIDTNGDGTPDATVTADPSGAW------TYTP-ST 747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              666 SSYPAVFVVGT----IGGVTGAYRS----DDCGTTWVLINDDQHQYGNWGQAITGDHANL 717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEA 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 597 IGTVTDDAGSVVGAIASGGSTDDATPTL-SGT----AEAGSTVSVYDGTTLLGTTTADPS 651
compound's activity; {11} a culture comprising strains in which the gene
               product is overexpressed or underexpressed, (12) determining the extent to which each of the strains is present in a culture or collection of strains, or (13) identifying the target of a compound that inhibits the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----TGPSASVTVTGDTTAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    173 ANGRWSYTPGTLPEGAHOITVTQTDSSGNTSELSTVGIVVDTIVQKTPAITAVSDDAQHP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPL
                                                                                                                                                                                                                                                                                                                                                                                                                                             ---IPLLDWVGWNNWGYNGVVSIAADPI--NTNKVWA------
                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TSLSGTAEAGATINIDTNGDGTPDAT-----VTADPGGAWTYTPSTPLPA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84 -AVGMYTNSWDPN-------DGAIL---RSSDQGATWQITPLPFKLG-
                                                                                                                                                                                                                                                                                                                                                                                                   Indels 243;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RDGGATWOAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGP---YDGS----
                                                                                                                                                                                                                                                                                                                                                            Length 2435;
                                                                                                                                                                                                                                                                                                                                                        4.9%; Score 199.5; DB 6;
23.0%; Pred. No. 0.00022;
live 72; Mismatches 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLPAGTVIGYTATDAAGN-----
                                                                                                                                                                                                                                                                                                                                                                                                   Matches 185; Conservative
                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                             47 ANGRW----
                                                                                                                                                                                                                                                                                                                   Sequence 2435 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             547
                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                909
$$$$$$$$$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
```

365

용

à

```
966 TYTVDTVAPNAPV-----LDPINATDPVSGQAEPGSTVTVTYPDGTTATVVAGPDGSW 1018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1295 VTATDAAGNAGTDTAVVTIDTTAPNA--PVLDPINATDPVSGTAEAGSTVTVTYPDGTTA 1352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     534 ADGSRFVWAPGDPGQPV----VYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGT 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Antisense; prokaryotic essential gene; cell proliferation; drug design.
                                                                                                                                                                                                                                                                                                                                                                                                                                    590 FYRSTDG--GVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA
                                                       749 ADKTIDAKVTFTDAAGNSSSVNDTQTYTIDTTAPDAPVINPVNGTDPITGTAEPGSTVTV
                                                                                                                                                                       ----GVQPNPPVPSPKLGWMDEAMAIDPFNSD-----RMLY--GTGATLYATND---
                                                                                                                                                                                                                                                                                                                                                                                                                -----LIKWDSGQIHI------APMVKGLBETAVND-----L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --TSVDYAELNPSIIVRAGSPDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                648 ITGVSSAVNVG------GKSAPGSSYPAVFVVGTIGGVTGAYRSDDC-----GTTW
                                                                                         SRDGGATWQAVPGAP-TGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ISPPSGAPLISA-----LGDLGGFTHAD-----VTAVPSTIFTSPVFTTG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1353 TVVAGTDGSWSVPNPGNLVDGDTVTATATDPAGNTSLPGTGTVSADI 1399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              694 VLI -- NDDQHQYGNWGQAITGDHANLRRVYIGTN ---- GRGIVYGDI
                                                                                                                                                                                                                                                     SWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF
                                                                                                                              TYPDGSTTTVVAGPDGTWTVPNPGLNDG-DKVTAIATDPAGNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protein encoded by Prokaryotic essential gene #2169.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABU16642 standard; protein; 1308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Acinetobacter baumannii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200277183-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21-MAR-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-OCT-2002
                  205
                                                                                                                                    809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABU16642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                        g
                                                                                                                                                                       δ
                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New Acinetobacter baumanii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    103 SDQGATWQITPLPFKLGGN--MPGRGMGERLAVDPNN-------DNILYFGAPSG 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VGWNNWGYNGV--VSIAADPINTNKVWAAVGMYTN-----SWD-PNDGAILRS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           V------NGTDP1TGTAEPGSTVTV-----TYPNGDTATVVAGPDGSWSVPNPG--LND 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GDEVEAIATDP----AGNPSLPGTATVD--AVGPNTDGVNFTVDSVTADNVINASEASG 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                K----GLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQ 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The invention relates to isolated Acinetobacter baumannii nucleic acids and polypeptides are useful as reagents accids and polypeptides are useful as reagents acciding a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii inflection, and as biocontrol agents for plants. The present sequence represents the amino acid sequence of an A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TWSNVAIGGGGFVDG-----IVFNEGAPGILYVR-TDIGGMYRWDAANGRWIPLLDW
                                                                                                                                                                                                                                                                                              Acinetobacter baumannii, bacterial disease; antibacterial; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90; Mismatches 344; Indels 309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3892;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.5%; Score 181.5; DB 21.5%; Pred. No. 0.0072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 5503; 328pp; English.
                                                                                                                                                                                                                                                       Acinetobacter baumannii protein #1377
740
                                      841
                                                                                                                                      Ź
718 RRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                    ADA34216 standard; protein; 3892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GENO-) GENOME THERAPEUTICS CORP.
                                    --VYDGTTLLGTTTAD---PSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0088701P
                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                           Acinetobacter baumannii
                                                                                                                                                                                                                                                                                                                  plant biocontrol agent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2003-576092/54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                G, Bush D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  baumannii protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 3892 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; ADA30090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-JUN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                JS6562958-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Simi
hes 204;
                                                                                                                                                                                                                  20-NOV-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                      13-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  543
                                    825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             plants.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Breton
                                                                                           RESULT 1
```

셤

⋧

셤

8

엄 8

8

727

491

576

```
TAP--VVALDDVLTNDSTPALTGTVNDPTATVVVN--VDGVDYPAVNNGDG-TWTLADNT 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
                                                                                                   233 TGTVNDPTATVVVNVDGVDYPAVNN------GD-GTWTLADNTL--PTLADG 275
                                                                                                                                                       PVFTTGTSVDYAEL--NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVT-- 525
                                                                                                                                                                                                                                                                                                                                                        577 VNPKTFYALSNGTFYRSTDG--GVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSG 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Antisense; prokaryotic essential gene; cell proliferation; drug design.
                                                                                                                                                                                    ---SVPNPGNLVDGDTVTATATDP-AGNTSLPGTGTVSADITAPVVALDDVLTNDSTPAL
                                                                                                                                                                                                                                                                                       331 TYPDGTTATVVAGTDGS---WSVPNPGNLVDGDTVTAT--ATDPAGNTSLPGTGTVSADI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              492 AVRLEPGSTVTVTYPDGTTATVVADTD---GSWSVPNPGNLVDGDTCDLSQSYXSLSGNT
                                                      GGQIH-----IAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTS
                                                                                                                                                                                                                                                          ----TGGTVAASADGSRFVWAPGDPGQPV----VYAVGFGNSWAASQGVPANAQIRSDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                           LYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          441 LPTLADGPHT-ITVTATDAAGNVGYD-----TAVVTIDTVAPKCTGAXSDQCNRLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -GTTWVLINDDQHQYGNW-----GQAITGDHANLRRVYIGTNGR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ohlsen Forsyth F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protein encoded by Prokaryotic essential gene #8219.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 25; SEQ ID NO 50616; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABU22692 standard; protein; 1074 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malone C,
Carr GJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-MAR-2002; 2002WO-US009107.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2002US-0362699P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ELIT-) ELITRA PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 549 SLPGTGTVSADI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GIVYGDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Burkholderia mallei.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; ACA26562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200277183-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-MAR-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 03-OCT-2002
     177
                                                                                                                                                       470
                                                                                                                                                                                                         276
                                                                                                                                                                                                                                                          526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABU22692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        oʻ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wang
Wall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABU22692
                                                                                                   유
                                                                                                                                                                                              q
                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                    The invention relates to an isolated nucleic acid comprising any one of the fall antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid nucleic acid, (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense unclaic acid; (4) an antibody capable of specifically binding the polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of agene in an operon required for proliferation, (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for publication or that inhibits proliferation, (8) identifying a gene capacity or that inhibits proliferation of an organism acts; (9) manufacturing an antibioic; (10) profiling a compound, activity; (11) a culture comprising strains in which the gene or which he each of the strains is present in a culture or collection of an organism. The antisense nucleic acids required converting or streaming for homologous nucleic acids required converting proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation to isolate candidate molecules for this or the terget prokaryotic essential genes. Note: The sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent in a contract of the printed specification, but was obtained contractive for proliferation to isolate capacitication, but was obtained contractive contractives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           136 VL-----DPINATDPVSGQAEPGSTVT-----VTYPDGTTATVVAGTDGSW----- 176
                                                                                                                                                                                                                                                                                    screening
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SSGDVWKFSVTSGTWTRI-----SPVPSTDTANDYFGYSG-----LTIDRQHPNTI 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --GD------GTWTLADNTLPTLADGPHTITVTATDAAGNVGNDTAVVTIDTVAPNAP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             369 PNPPVPSPKL---GWMDEAMAIDPFNSDRMLYGTG-----ATLYATNDLTKWDS---- 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIOGVVWVAFDKSSSSLGQASKTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GSWSVPNPG-----NLVDGDTVTATATDPAG------NTSLPGTGTVS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PVGVADPNNPVFWSRDGGATWQAVPGAPTGFI - - PHKGVFDPVNHVLYIATSNTGGPYDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          314 MVATQISWWPDTI----IFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----ADITAPVV-ALDDVLTNDSTP-ALTGTVNDPTATVVVNVDGVDYPAVNN----
                                                                                                                        Zyskind JW;
Xu HH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                       New antisense nucleic acids, useful for identifying proteins or scre
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.5%; Score 181; DB 6; Length 1308;
24.1%; Pred. No. 0.0018;
ative 65; Mismatches 245; Indels 200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          <del>::</del>
-:
                                                                                                                        Ohlsen KL,
Forsyth RA,
                                                                                                                          Haselbeck R,
                                                                                                                                                    Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 44566; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       directly from WIPO at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ftp.wipo.int/pub/published pct sequences
                                                                                                                          Malone C,
Carr GJ,
08-FEB-2002; 2002US-00072851.
                                                                                                                          Malone
                                                                      (ELIT-) ELITRA PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                          Zamudio C,
Trawick JD,
                                                                                                                                                                                                   WPI; 2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 162; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1308 AA;
                                                                                                                                                                                                                             N-PSDB; ACA20512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          electronic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86
                                                                                                                          Wang L,
Wall D,
```

Zyskind JW; Xu HH;

ŘŘ

g

ઠે g ò g ઠે g

proliferation, (1) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibicit; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the confiferation of an organism. The antisense nucleic acids required for callular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for ftp.wipo.int/pub/published_pct_sequences

Sequence 1074 AA;

Query Match 4.2%; Score 170; DB 6; Length 1074; Best Local Similarity 21.6%; Pred. No. 0.008; Matches 186; Conservative 82; Mismatches 300; Indels 292; Gaps Best Local Similarity 21.6 Matches 186; Conservative

g

ò d

ò

ò

ò g à g ò g 용

ð 셤

ઠે

	522 GGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIR 573	639 SAPTIPNGGKPSQDLTLVGAASGPVALHNVA-PGTASTDAVNVGQLGAVTTGLGGGAAI- 696	GVTFC	697DPKTGAVTAPSYTVYNADGTTSNVSNVGAAIDAINSTGIKYFHANSTKPDSQALGA 752	615VMPHAVPGKEGDLMLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGK 661	753 DSVAIGPNAVANNAGDVALGSGAVTSQAGGTLSETINGVTYSFAGTTPIGTVSVG 807	662 SAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWG-QAITGDHANLRRV 720	808 -APGVERTITNVAAGRIGOSSTDAINGSOLYGTNOSIEALTDKANSLGNT 856	721 YIGTNGRGIVYGDIGGAPSG 740	857 VANTLGSGASYNPQTGAVNG 876
Ωp	ò	qq	ò	q	ò	qq	ò	qq	ò	QQ

Search completed: May 11, 2004, 12:06:49 Job time : 125.24 secs

```
version 5.1.6
- 2004 Compugen Ltd.
GenCore
Copyright (c) 1993
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 11:51:45; Search time 34.7001 Seconds (without alignments) 2051.340 Million cell updates/sec

US-09-917-376-3 4036 1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283366 seqs, 96191526 residues Searched:

Total number of hits satisfying chosen parameters:

seq length: 0 seq length: 200000000 Minimum DB Maximum DB

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	probable secreted		Avicelase III - As	hypothetical prote			_	hypothetical prote	tail-host specific	probable PPE prote	outer membrane sec		probable RTX famil	hypothetical prote	proteoglycan core	extracellular seri	hypothetical prote	OI	hypothetical prote	nucleoporin - rat	probable structura	related to stress	ical p	8	probable PPE prote	hypothetical prote	;;	1.)	_
SUMMAKIES		T35237	D97013	T00349	F72393	A83412	E95965	875251	S76412	T13256	F70846	AE1905	AE2254	B85547	F90696	A28452	AB3528	H87323	H90681	D70533	A44345	D85532	T49741	T16580	S10789	E70969	H98323	8	68	T34434
	DB	2	~	Ŋ	(7	7	7	~	7	~	~	N	0	N	~	~	N	7	~	N	~	N	~	7	~	N	N	0	N	7
	Length	1	839				2174	3972	4199	1904	2523	993	908	5188	5291	2124	2554	902	980	669	1468	980	618	13055	1684	3716	1341	62	44	2232
d	Query Match	58.7	49.8	41.6	28.8	5.9	5.0		4.6	4.5	4.3	4.3	4.2	4.2	4.2	•	•	4.1	4.1	4.1	4.1	•	4.1	•	4.0	•	4.0		3.9	
	Score	2371	2009	1680	1160.5	238	202.5	189	186.5	182	175.5	172	171	170.5	168.5	168	167.5	166.5	166.5	166	166	165.5	163.5	163.5	163	160.5	160	160	158.5	158
	Result No.	1	2	m	4	<u>.</u>	9	7	8	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

quinate-shikimate	probable PPE prote	probable PPE prote	hypothetical prote	hypothetical prote	mucin MUCSB, trach	OmpA-related prote	hypothetical prote	mucin, submaxillar	hypothetical prote	probable oxidoredu	beta transducin-li	beta-amylase (EC 3	hypothetical 367K	hypothetical prote	hypothetical prote
A55547	A70524	B70969	G83559	T34433	T45025	A87364	AG2335	T03099	877300	T35248	T42045	A29130	T31308	B72575	B75625
7	~	N	~	7	~	7	~	~	7	7	~	~	~	7	8
809	2204	3157	4180	1032	3570	1055	1821	13288	3016	348	1049	1196	3472	820	1145
3.9	3.9	9.6	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8
		99	55	2	5.	153	153	153	2.5	152	152	152	152	1.5	51.5
157	156.5	ī	7	154	154		•		15					15	151

ALIGNMENTS

RESULT 1 T35237 Probable secreted cellulase - Streptomyces coelicolor C;Species: Streptomyces coelicolor C;Gpecies: Streptomyces coelicolor C;Gpecies: Streptomyces coelicolor C;Date: 05-Nov-1999 #text_change 05-Nov-1999 C;Cacession: T35237 R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. Submitted to the EMBL Data Library, September 1998 A;Reference number: Z21572 A;Reference number: Z21572 A;Residues: preliminary; translated from GB/EMBL/DDBJ A;Residues: 1-890 <see> A;Residues: 1-890 <see> A;Residues: 1-890 <see> A;Cross-references: EMBL;AL031515; PIDN:CAA20642.1; GSPDB:GN00070; SCOEDB:SC5C7.3 A;Genetics: C;Genetics: A;Genetics:</see></see></see>

30c

10; 66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125 65 6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY Gaps 24; Query Match Best Local Similarity 57.7%; Pred. No. 1.4e-131; Matches 431; Conservative 109; Mismatches 183; Indels 셤 ò

160 126 GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS 185 220 244 101 TGVVALASDAVDPDRVYAAVGTYTNDWDPTNGAVLRSADRGASWBKADLPFKLGGNMPGR 161 GMGERLAVDPHDNDVLYLGAPSGHGLWRSTDAGVTWSEVTAFPNPGNYAQDPNDTSGYAS 186 DIQGVVWVAFDKSS-SSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK g ò 셤 ઠે

304 245 GVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLT

> g ठे

281 GVLDAENGYLYLAYSDTGGPYDGGKGRLYRYATATGTWTDISPAAEADT---YYGFSGLT 337 364 305 IDROHPNTIMVATOISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLT a ò 셤

ò g

VKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAEL 483 424

```
hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PAO1)
hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Species: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: A83412
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Briadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Nature 406, 959-964, 2000
A;Fitle: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A,Molecule type: DNA_A,Residues: 1-2468 <STO>A,Residues: 1-2468 <STO>A,FOross-references: GB:AE004613; GB:AE004091; NID:g9947856; PIDN:AAG05263.1; GSPDB:GN001: A,FXperimental source: strain PAO1 C,Genetics: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERPERICE: A,FERP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1084 VNLSNGSSLSG----TAEPGSTVILTD------GNGNPIAEVTADGSGNWTYTPSTP 1130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1188 -DAGGNPIGQVTAD-----GTPLA 1214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116 FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIA 175
                                                                                                                                                                                                                                                                                                                                            123 MVK-----GLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGT 476
                                                                                                                                                                                                                                                                                                                                                                           497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      303 LTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPW 362
                                                                                                                                                                                                                   363 LTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAP 422
240 LRGKIH--DGILYVTLSNALGPNGATRGAVMKYVIADQKMYDVTFMKGD-----FGYCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  498 NDGKTLVWSPAN--HEVIVSSDKGKSWKKAISVPVPEFNYFPASDPVNPSKFYIFDWKNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    556 DFLISKDGGKSFMK-GAKLPSFDNWWVSLYSFPVLAPDREGDIWLALQWNGLYRSKDGGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     340 IK---DLNP------HWISD-VKIDPFDMNRAIFTTGYGVWVTYELKKSFEG----MGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          444 GIAFAYONSKFVARVHTY---TYP---FLSYSEDGGINWREIETVPEGITDGGRLSLAVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --NG-VVSIAADPINTNKVWAAVGMYTNS-----WDPNDGAILR-SSDQGATWQITPLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGG--TVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         534 ADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA---NAQIRSDRVNPKTFYAL--SNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              589 TFYRSTDGGVTFOPVAAGLPSSGAVGVMFHA----VPGKEGDLWLAAS-SGLYHSTNGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11 VAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 2468;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.9%; Score 238; DB 2; Length 24 22.8%; Pred. No. 9.1e-06; tive 88; Mismatches 314; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 575 G-WIHYMIGDMNEFGRIFLGTEGRGIIVGEV 704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 22.8%
Matches 205; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A83412
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                       g
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72393
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary
A;Molecule type: DNA
A;Molecule type: DNA
A;Residues: 1-77 cARNA
A;Cross-references: GB:AE001712; GB:AE000512; NID:g4980799; PIDN:AAD35393.1; PID:g498080
A;Experimental source: strain MSB8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A,Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome A,Reference number: A72200; MUID:99287316; PMID:10360571
A,Accession: F72393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26;
                                                                                                                                                                                                                                                                                                                                                                                                                                    594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GWSFGFGKASSTGSYVVIYGFFTIDGAAGLFKSEDAGTNWQVISDASHGFGS-GSANVVN 717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414
                                                                                                                                                                                       474
                                                                                                                                                                                                                                                                                                                                                                           543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            595 DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLP---FKLGGNM 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 MGVLSVALDPSDPKRIYAMTGKYTQDW-AGYGAILISEDYGETWTIVNLDKYGIKVGGNE 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSG 302
      TYYGYGGLSVDLQVPGTLMVAALNCWWPDELIFRSTDSGATWSPIWEWNGYPSINYYYSY 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 80
                                                                                              535 DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                       415 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                               GISVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183 YQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIP
                                                                  DISAEPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 28.8%; Score 1160.5; DB 2; Length 707; Best Local Similarity 35.2%; Pred. No. 1.2e-60; Matches 264; Conservative 123; Mismatches 275; Indels 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hypothetical protein TM0305 - Thermotoga maritima (strain MSB8)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nature 399, 323-329, 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C,Genetics:
A,Gene: TM0305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123
      312
                                                                                                                                                                                                                                                                                                               475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    જે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      유성
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                               ઠે
                                                                                                                        셤
                                                                                                                                                                    ò
                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
```

703

50;

36;

Length 2174;

DB 2;

90; Mismatches 312;

------TWOITPLPFK 117

```
1012 GGASGGN-----AIVGL-----DSGTHLQTSGSGARGVIVQSIGGGGGTSQGASVGLS 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | | | : | : | : | 33.2
| SB3_IGGGGG----QGGDGSDEAAGRITVGGGFGGSGGAGGSGGWVTAKGWINLSTSG----- 1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1437 GAVTVDLTSQSSI---TTQGTAAHGLVAQSIGGGGGVGGAASGAPLSFTGNSPGSYGDGG 1493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1494 DVAVTADGSIFTRGDYAFGVLAQSIGGGGGFGGN-ATSAFIGSNGNLSSDG----- 1543
                                                                                                                                                                                                                                              060 ASASLPGGG-EEAAEAEESEEGSGAFSASVGVSVGRTGGSGGS-SGTVNVTTAGT---- 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----DDAHGIVAQSIGGGGGVGGAGSSTAAEKEHSHTIDLVVGGSGG----VGGSGGE 1382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GATWQAVPGAPTGFIPHKGVFDPVNHVLYIATS------NTGGPYDGSSG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRS 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TDGGATWTRIWDWTSYPNRSLRYVLD-------ISAEPWLTFGVQPNPPVPS 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EETAVNDLISPPSGAPLISAL-----GDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               376 PKLGWMDEAMAI - DPFNSDRMLYGTGATLYA - - - - TNDLTKWDSGGQIHIAPMVKGL 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      226 GISTASGSQATANITVGVGGSGGAGGAGGAVGITFDDNHGNSISTAGYSAY---GVLLQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --ISTFGADADGVLAQSIGGGGGGGGGVGQASSGDSEPLDDEGDSECASEGGNGDDGHGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TVAASADGSRFV---WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYAL
                                                                                                                                                                                                                12 AIGGGGFVDGIVFNEGAPGILYVRT-DIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVS
                                                                                                                                                                                                                                                                                                                                                                                                         LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TDTTGYQSDIQGVVWVAFDKS---SSSLGQASKTIFVGVADPNNPVFWSRDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AELNPSIIVRAGSFDPSSQPNDRH--VAFSTDGGKN------WFQGSEPGGVTTGG
                                                                                                                                                                                                                                                                                                             71 IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA-
                                                                                                                       5.0%; Score 202.5; DB 22.7%; Pred. No. 0.00093
                                                                                                               Query Match 5.0%
Best Local Similarity 22.7%
Matches 176; Conservative
    A; Contents: annotation
                     C.Genetics:
A.Gene: SMb21548
A.Genome: plasmid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1604
                                                                                                                                                                                                                                                                                                                                                                                                         118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  875251
                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: Sinorhizobium meliloti
C; Species: Sinorhizobium meliloti
C; Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C; Accession: E95965
R; Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A; Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endc A; Reference number: A95842; MUID:21396508; PMID:11481431
A; Accession: E95965
A; Molecule type: DNA
A; Status: preliminary
A; Molecule type: DNA
A; DATA A; Arm.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Residues: 1-2174 KURN-
A;Residues: 1-2174 KURN-
A;Cross-references: GB:AL591985; PIDN:CAC49389.1; PID:g15140875; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
A;Experimental source: strain 1021, megaplasmid pSymB
B;Galibert, F:; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, pela, D.; Chain, P.; Cowie, A.; Poros, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  maga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical glycine-rich protein [imported] - Sinorhizobium meliloti (strain 1021) C, Species: Sinorhizobium meliloti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TE-ISGTAEAGATVILT-----DGGGN----PIGQATADGSGNWTFTPSTPLANGTVIN 1548
                                                                                                                                                                                                                     SLPATTÍVDSSLPSIPQVDPSNGSVISGTADAGNÍII--ÍDGNGNPIGQVTADGSGNWS 1452
                                                                                                                                                                                                                                                                                                                                                                                                                          1453 FTPGIPLPDGTVVNVVARSPSNVDSAPAVITVDGVAPAAPV-------IDPSNG 1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----LSNGTVVNAVAQDAAGNTSGPVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVI 1759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LTDGNGNPIGQTLADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTSGPASTTVDTVAPAT 1819
                                                                                                                                     1275 PIGETTADGSGNWSFTPGTPLANGTVVNAVAQDPAGNTGPQGSTTVDAVAPNTPVVNPSN 1334
                                                                                              251
                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SALGDLGGFTH-----ADVTAVPSTIF--TSPVFTTGTSVDYAELNPSIIVRAGSFDPS 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----- QGSEPGGVT----- 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----TGGTVAASADGSRFVWAPGDPGQPV--VYAVGFGNSWAASQGVPANAQIRSDRVNP 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KTFYALSNGTFYR-----STDGGV--TFQPVAAGLP---SSGAVGVMFHAVPGKEGD
  ---PIDTTGYQS----DIQGVVWVAFDKSSSSL--GQASKTIFVGVADPN-N
                                                                                                                                                                                      HVLYIATSNTGGPY - - - DGSSGDVWKFSVT - SGTW - - TRISPVPS - - - - TDTANDYFGYS
                                                                                                                                                                                                                                                                                GL----TIDRQHPNTIMV----ATQISWWPD---TIIFRSTDGG------ATWTRIWD
                                                                                                                                                                                                                                                                                                                                                                              WT----SYPNRSLRYVLDISAEPWL--TFGVQPNPPVPSPKLGWMDEAMAIDPFNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        393 DRMLYGT---GATLYATNDLTKWDSGGOIHIAPMVKGLEETAVNDLISP----PSGAPLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GO------VTADGSGNWSFTPGTPLANGSVINALAQDAAGNNSSPTSATVDSLAPAAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        658 IDPSNGSVIAGTAEAGATVILTDGNGNPIGQVTADGSGN-WSFTPGTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LWLAASSGLYHS-TNGGSSWSAITGV----SSAVNVGFGKSAPGSSYPAVFVVGTIGGVT
                                              1215 NGTVIVATAIDPTGNTGPQAATTVDAVAPPAPVIDPSNGTTISGTAEAGAKVILIDGNGN
                                                                                            PV-FWSRDGGATWQAVPGAP------TG-----FIPHKGVFDPVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------PVINPSN-----GSVITGT-AEVGAKVILTDGNGNPIGETTADGSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOPNDRHVAFSTDGGKNWF ----
                                                                                                                                                                                                                                    1335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1549
                                                                                            219
                                                                                                                                                                                                                                                                                  302
                                                                                                                                                                                                                                                                                                                                1395
                                                                                                                                                                                                                                                                                                                                                                              343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                         셤
                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                   ò
                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
ઠે
                                                                                       ઠ
```

```
Aypothetical protein slr1028 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Arziety: PCC 6803
C;Accession: S75251
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; R;Kaneko, T.; Sato, S.; Ehimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.
                                                                                                                                                                                                                                                                                                                                                     *;
                                             SNGTFYRSTDGGVTFQPVA-----AGLPSSGAVGVMFH-AVPGKEGD----LWL
                                                                                                                                                                                                         SAGKDNVVTVNSGGNVSAASGV--AVQYTAGMNSPEDSTLVVNNAGTISGSVKGA 1656
                                                                                                                                                            AASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTI-GGVTGA
```

```
hypothetical protein slr0408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      459 VTAVPSTIFTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S:
Matches 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
2515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3015 LAVLAPGYRNILQFPILDY-----PAINNLGGVFIYYGEESGISVNDPPDVVLAAPDL 3067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2812 YVFFVKDASDNDILYSSSSNPG----SSSG--WDGTSTVLTFSDVNQATN-----FPL 2858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SATVVPGLDGDTLAVA-----FRSNNSPATWVGLLNSSDVTNWQGSAELTQVDA 2907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2908 NSQVSLTVVDGTYYLFFTSSTEASASYATSTDGLNWGDITLIPW--------DDGNL 2956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PQETIFELLEITFTGDVNGDGFDDLLISAPLTPVIAGQFPDVNGDQGVSWVVFGGTHWGT 3127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EYTANSPFGLGNLANNQTNNSQNFNPYGFVTTGLPRSQAGISISGGADVNGDGFSDFALG 3187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3188 APGNFDNLSYVLFGSDFTNQVNQLGTIGDDVMLGSPTGEIFVAGQGDDQIYTNGGVDTVY 3247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2753 PSITAYNGTLYLGFTADNGQV-WVAEGVNTNANSGILINAVPINQASNNGPTLVAFNDEL 2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             342
                                                                                                                                                                                                                                                                                                                   42,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TTGTSVDYAELNPSIIVRAGSFDP--SSQP------NDRHVAFSTDGGKNW-- 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----GGVTTG-----GTVAASADG-SRF-VW 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---PGOPVVYAVGFGNSWAASOG---VPANAQIRSDR 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GVF----DPVNHVLY1ATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY 300
                                                                                                                                                                                                                                                                                                                                                                   57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VLQGPNSTIAVYGNANGTLTDSSPLTF--GNNFPLPSSFTGIDLNGNGIKEIVAGOPNLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDGLADFGVIDDQGNFFLVLGSPELGSQGSLVIDSTLPNLSNFNQAWGVGDFNGNGYDDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | : | | : | |: | |: | |: | |: | |: || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || : |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SFATLDGWLYQAFYGINERISTKDSYIYIQRSRDGVSWENLTQVVPLDSNGTPIDLKNLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -SLGQASKTIFVGVADPNNPVFWSRDG-----GATWQAVP-----GAPT--GFIPHK
                                                                                                                                                                                                                                                                                                                                                                13 IGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANG-------RWIPLLDW-
                                                                                                                                                                                                                                                                                                                                                                                                                   VNGDGFADVISGGSPAGGVLIFGNSTKDLL--DAALGTDDLIISVENAQVKEFVALGDFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -------VGMNNWGYNG---VVSIAADPINTNKVWAAVGMYTNSWD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -------GRGMGERLAVDPN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----NDNILYFGAPSGKGLWRST-----DSG-ATWSQMTNFPDVGTYIANP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --OSDIQGVVW-----VAFDKSSS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRI------WD------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----WISYPNRSLRYVL-----DISAEPWLTFGVQPNPPVPSPKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GWMDE-----AMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---IAPMVKGLEETAVNDLISPPSGAPLISALGDL----GGFTHADVTAVPSTIFTSPVF
                                                                                                                                                                                                                                                             4.7%; Score 189; DB 2; Length 3972;
19.3%; Pred. No. 0.013;
tive 93; Mismatches 276; Indels 326; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----PNDGAILRSSDQGATWQITPLPFKLGGNMP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3248 AGPGNDFVTVTDTNFRRLDGG 3268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    577 VNPKTFYALSNGTFYRSTDGG 597
                                                                                                                                                                                                                                                                                          Best Local Similarity 19.3
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TDTT--GY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -FOGSEP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             542 APGD---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3128
                                                                                                                                                                                                                                                                                                                                                                                                                         2458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 473
                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
```

```
Cispecies: Synechocystis sp.
A;Variety: PCC 6803
C;Date: S-Apr.1997 #sequence_revision 25-Apr.1997 #text_change 20-Jun-2000
C;Accession: S76412
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule fype: DNA
A; Residues: 1-4199 «KAN»
A; Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BAA18541.1; PID:g165362:
A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C; Superfamily: Synechocystis hypothetical protein slr0408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2318 FTVSTNNGOSWNSEVQL---PQSAQTIFPPALAFFNNVLYLAYVDGNNGLNIITSQDQGQ 2374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2544 ANNELYLS-----YPGPQGS-QELNVTS-----FPLPFTGSILGNGSLVRFLGDVNG 2589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2809 LAVLASNYYAAAIGEPNNLPNYLSRPGNQGGVFIFYGNSNGLSNTAQPDVILAAPFTNPS 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2432 NOTAISAISATVLGDTLYLVYKGGTRNTPSTLDYITSTTN-----ADLSANDWSSIPIPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -- DKSSSSLGOASKTIFVGVADPNNPV-FWSRDGGATWQAVPGAPTGFI---PHKGVFDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---DIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DYFG--YSGLTIDRQHPNTIM-----VATQISWWPDTIIFRSTDGGATWTRIWDWTSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---OPNPPVPSPKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TWQITPLPFKLGGNMPG-----RG-MGERLAVDPNNDNILYFGAPSGKGLWRSTDSGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TWN---APLALGGTSSTPPTLFVYQGTLSLLFAANNSTSTVLQFYLNSSNEWIYANEIGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----SGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PVFTTGTSVDYAELN-----PSIIVRAGSFDPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DWVGWNN-WGYNGVVSIAADPINTNKVW-AAVGMYTNSW-----DPNDGA-ILRSSDQGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9 SNVAIGGGGFVDGIV------FNEGAPGILYVRTDIGGMYRWDAANGRWIPLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                160 TWSQMT-----NPPDVGTYIANPTDTTGYQSDIQGVVWVAF---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2487 VSSQGGPSLTNDGTNLYLSYLDSSNQLNFVSSGNGINWSS-PQVITNNISQSPPAIAF--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----PNDRHVAFSTDGGKNW-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 389;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 4199;
- Synechocystis sp. (strain PCC 6803)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GWMDEAMAIDPFNSDR----MLYGT----GATLYATNDLTKWD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     250 VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST--
                                                                                                                                                                                                                                                                                                    A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76412
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Autch 4.6%; Score 186.5; DB 2; Local Similarity 19.2%; Pred. No. 0.019; les 198; Conservative 124; Mismatches 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PNRSLRYVL---DISAEPWLTFGV------
```

Db 1119 PNNGWTSTVP-TVAEGNYLWTKTVWTYTDNTS 114	369 PNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM	bb 1150ETGYSVAMMGVKGDKGDPGNNGTNGIAGKDG 1180	QY 424 VKGLBETAVNDLISP-PSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFT 473	OY 474 TGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWF 516	OY 517 QGSEPGGVTTGGTVAASADGSRFVWAPGDPQQPVYYAVG 555	Qy 556 FGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAG 606 :	Qy 607LPBSGGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSS 644	Qy 645	RESULT 10 F70846 probable PPE protein - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Date: 17-U1-1998 #sequence_revision 17-Ju1-1998 #text_change 22-Oct-1999 C;Accession: F70846	R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998 A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.	A;Title: Deciphering the biology of Mycobacterium tuberculosis from the comple A;Reference number: A70500; MUID:98295987; PMID:9634230 A;Accession: F70846 A;Status: preliminary; nucleic acid sequence not shown; translation not shown A;Molecule type: DNA	A;Residues: 1-2523 <ccd.> A;Residues: 1-2523 <ccd.> A;Experimental source: strain H37Rv C;Genetics: A;Gene: PPB</ccd.></ccd.>	Query Match Query Match Best Local Similarity 21.8%; Pred. No. 0.044; Matches 175; Conservative 92; Mismatches 327; Indels 209; Gaps 38; Ov 10 NVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAAN-GRWIDILDWVGWNNWGYN 66	1522 NTGIGNSGIASTGLENAGGENTGVNAGGSVNTGSENAGGANTGGENPGSVNTGMLNTG 67 GVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSS-DQGATWQITPLPFKLGGNMPG
2869 GQISTYQLSRIAQAGDVNGDGFDDLLISSPYTVDAENNQGGVFVVFGGDDWNNQPFDLGQ 2928		2929 LRANQSGGSNPRGFAIDGSFNSQAGIALNGGGDINGDGFADFIIGAPGENNLQYNQQIVF 2988	545DPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYAL 585 2989 IENGELSDDDKXSYILXLDGNQTIQMGGGDWQANQ-VWTN-QVATNWNNSSRPPEAVIGQ 3046	586 SNGTFYRSTDGGVTFQ	624 EGDLWLAASSGLYHSTNGGSSW-SAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTG 682 : :	683 AYRSDDGGTTWVLINDDQHQYGNWGQAITG	726 GRGIVYGDIG 735 	T13.2-Nost specificity protein homolog - Lactococcus lactis phage BK5-T C;Species: Lactococcus lactis phage BK5-T C;Species: Lactococcus lactis phage BK5-T C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999 (C;Accession: T13256 Bs.; Hillier, A.J. Appl. Environ. Microbiol. 61, 4089-4088, 1995 A;Tille: Sequence analysis of the temperate Lactococcus lactis bacteriophage BK5-T A;Reference number: Z17646; MUID:96064422; PMID:8526523	# type: DNA 1.1904 <box> 1.1-1904 <box> 2. 1-1904 <box 1-1904<="" 2.="" td=""><td>; Conservat ATTQPYTWSNVA</td><td>WIPLLDWYGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA : </td><td>108 TWQITPLPFKLGGMMPGRGMCERLAVDPNNDNILFFGAPSKGLWRFTDSGATW-SOWIN 166 1: : </td><td>974 VPAGQFLWTKTVWTYTDN-TSETGYSVAMYGVKGDKGDPGINGTNGIA 1020 219 PVFWSRDGGATWQAVPGAPTGFIPHKGVFDP 249 : </td><td></td></box></box></box></box></box></box></box>	; Conservat ATTQPYTWSNVA	WIPLLDWYGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA :	108 TWQITPLPFKLGGMMPGRGMCERLAVDPNNDNILFFGAPSKGLWRFTDSGATW-SOWIN 166 1: :	974 VPAGQFLWTKTVWTYTDN-TSETGYSVAMYGVKGDKGDPGINGTNGIA 1020 219 PVFWSRDGGATWQAVPGAPTGFIPHKGVFDP 249 :	

```
hypothetical protein alr3588 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. pCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2254
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
                   45
                                                                                                                                116
                                                                                                                                                                                                                                                                                                                    234
                                                                                                                                                                                                                                                                                                                                                                                                                                338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AE2254
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            dd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                            8
                                                                                                                        g
                                                                                                                                                                 ò
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, B
DNA Res. B, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        outer membrane secretion protein alr0791 [imported] - Nostoc sp. (strain PCC 7120) C.Species Nostoc sp. PCC 7120 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120 C.bate: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2117
                                                                                                           PERSANVAIDSFIVPSIPIPQIDLAATIVSVGLGPITVPHLDIPRVPVTLNYLFGSQ--- 1709
                                                                                                                                                                                                                                                     LLASITGHSEPVDIFPGGLTIPAMNPLSINLSGGTGAVTIPAITIGEIPFDLVAHSTLG- 1827
                                                                                                                                                                                                                                                                                                                        1917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NTGVVNAGSYNTGSFNAGQAN--TGGFNPGSVNTGWLNTGDINTGVANSGDVNTGAFISG 2175
                                                                                                                                                                                  1710 -PGGPLKIGPITGLFNTPIGLTPLALSQIVIGASSSQGTITAFLANLPFSTPVVTIDEIP 1768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2016 NIGFGNAGSYNFGL-ANM-----GVGNIGFANTGSGNFGIGLTGDNLTGFGGFNTGSGNV 2069
                                                                                                                                                 292
                                                                                                                                                                                                                          334
                                                                                                                                                                                                                                                                                                   383
                                                                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                                                                                                   500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Cross-references: GB:BA000019; PIDN:BAB72748.1; PID:g17130136; GSPDB:GN00179
A;Experimental source: strain PCC 7120
 RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                 -----GLFNSGTGNVGFFNSGTGN----WGVFNSGSY-NTGIGNSGIASTGLFNAGGF
                                    -----TPFLDLTLGGLGSVVIPAIDIPAIR
                                                                                                                                                 VPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTD
                                                                                                                                                                                                                          TANDYFGYS-----GLTIDROHPNTI-------MVATQISWWPDTIIFRSTDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AP-LISALGDLGGFTHADVTAVPST1FTSPVFTTGTSVDYAELNPS11VRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          501 NDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV-WAP---GD----PGQPVVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --GNVGLGNVGDFNLGAGNVGGFNVGGGNIGGNNVGLGNVGWGNFGLGNSGLTPGLMGLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTF------YRSTDGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          599 TPQPVAAGLPSSGAVGV-MFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA-----ITG
                                                                        SDIQGVVWV-AFDKSSSSLGQ---ASKTIFVGVAD-----PNNPVFWSRDGGATWQA
                                                                                                                                                                                                                                                                                                 ATWIRIWDWISYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDE-----
                                                                                                                                                                                                                                                                                                                                                                          -AMAIDPFN-SDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 4.3%; Score 172; DB 2; Length 993;
Best Local Similarity 19.6%; Pred. No. 0.02;
Matches 168; Conservative 96; Mismatches 325; Indels 270; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :::| | | ::|
NYSN-----GAFWRGDYQGLLG 2192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DHANLRRVYIGTNGRGIVYGDIG 735
                                    LGFSYRPAVLPQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Residues: 1-993 <KUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: AE1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genetics:
Gene: alr0791
 125
                                    1615
                                                                        185
                                                                                                             1653
                                                                                                                                                 233
                                                                                                                                                                                                                                                           1769
                                                                                                                                                                                                                                                                                                                                    1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            713
                                                                                                                                                                                                                                                                                                                                                                                                                                                   442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11
ò
                                    유
                                                                        ò
                                                                                                     q
                                                                                                                                               ò
                                                                                                                                                                                  g
                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

44;

```
124
                                                                                                                                                                  DOTTCTSGSTANILAGKGGTTAFSTNYEGASITVTPVNDTAPTLNAGSPTLPTINEDAPL 175
                                                                                                                                                                                                                                                                     279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    291 TDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTI---------IFRSTDGG 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   335 ATWIRIWDWISYPNRSIRYVLDISAEPWLIFGVQPNPPV--------PSPK 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---WD-----SGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVP 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TGNDTLDCGEGSDLLYGNEDNDIINGGVGNDNLDG----GTGDD----ILRGGTGNDIY 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             639 TVDTVGDVIEENPNEGTDKVNSYISWTLGANLENLTLLGNTIIDG--TGNELDNHIIGNN 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AVGVMFHAVPGKEGDLWLAASSG--LYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYP 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---GOAITGDHAN 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -vyeidsvgdv--iieaadagidtvissvd-----WilgvnlenlilvgnQatlGignD 795
92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :| ||:
-SITVSPINNPIQGTTGLDKLYGTANEDIINGNEGNDYLFGRAGNDTLDGGEGNDYLFGG
                                                                                                                                                                                                                                                                                                                                             IANPIDITGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PGA---PTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        280 TGGIQLNSSSIGSSG-YSNYNSYAPILFNQAFPELDPVKGFTISFDVKINGETHTSD-DN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GDGIQDRAGFSVIVVTSDKTKAI----ELGFWTDEIWAQTASPLFTHSTTERAFRNTTTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VT--------RYHLVVENNTYKLFAPDSSTPILSGNLRDYSAFNHSTAAPSP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STIFTSPV-----FTTGTSVDYAELNPSII------VRAGSFDPSSQPNDRHVAFS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T----DGGK--NWFQGSEPGGVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---AASQGVPANAQIRSDRVNPKTFYA----LSNGTFYRST--DGGVTFQPVAAGLPSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   236 LLLPSIR----PVGGTQSLGSLGSLPTSQGWLKFGASPPVFPLF---PVGGTQSLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AILRSSDQGATWQITPL----PFKLGGN--MPG-----
                                                                                                                                                                                                                                 -RGMGERLA-------VDPNNDNILYFGAPSGKGLWR-STDSGATWSQMTNFPDVGTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNW----
DAANGRW-IPLLDWVGWNNWGYNGVVSIAADPINTN-KV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  378 LGWMDEAMAIDPFNSDRMLY-GTGAT-LYATNDLTK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LRRVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LDNRITGNNADNVLFGEAG 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | | : | |
494 FRAWDGSNGVAGG-----
```

```
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 191; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            - ტ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 523
                                                                                                                                                                                                                                                                                                                                                   A; Gene: Z0615
                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    substrain
                                                                                                                                                                                                                                                                                                                                                   41;
  A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-908 <KUR>
A;Residues: 1-908 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB75287.1; PID:917132721; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SET STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE STATE ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDG--GVTFQPVAAGLPSSGAVGVMFHA 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VVG--NTLYFTANNG----TSGHEIWQYQNGTTASLLEDITPGNNSFAP-SSLTAV---- 745
                                                                                                                                                                                                                                                                                                                                                                                                    IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQG---ATWQITFLPFKLGGNMPGRGM 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                            171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVP---GAPTGFIPHK 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OLI-----SSSNSVIOTSATV-----GTSLESISRSLAAGTYYIRAY 245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GVFDPVNHVLYI------ATSNTGGPYD------GSSGDVWKFSVT 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SQSDPGNYNLEFNFSADPPDAGGNTSDTSTPINLPATFSEIISDQVSLGDSSDYYQFTLA 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | : | | | | | | | | SASLVEIQFTSLTADANLY-----LQTQNGGNILSSTQPGTALDAVRLSLNAGTYNIL 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VNRGSTETAQYTLSGFAQAIGNDQAPNSTTLALNLNLNSPISLNEFVGNIDTNDYYKFTV 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                415 -GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFT 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TGTSVDYAEL---NPSIIVR---AGSFDPSSQ---PNDRHVAFSTD--GGKNW-FQG--- 518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ANDGINGVQLWSSNGDITTRLSNISSFNPTNLIVFNNRLYFAASNDTFGRELWEYNGTTV 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSA----VNVGFGKSAPGSSYPAVFVVG 675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----RYVLDISAEPWLTFGV 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEPGGVTTGGT----VAASADGSRFVWA-PGDPGOPVVYAVGFGNSWA 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTFTTTFNNQL-----FFTAQNNSQLWSTDGTIGGTQVISAGGITKSTPRNL--T 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GERLAVDPINIDILLYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QPNPPVPSPKLGWMD----EAMAIDP--FNSDRMLYGTGATLYATNDLTKWDS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NRISDINVGAGNSNPGNLTVAGNKLFFTAVDNDSIRKLWVYNGTNVNLVDVNASFSNSST
                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TIGGVTGAYRSDDCGTTWVLINDDQHQYG--NW---GQAITGDHANLRRVYIGTNGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          - - - - GNTLYFVTDSDNDFNLELWKSDGTAAGTD-----IIGTDGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            -------NDNTIANARNIGTFNTTTSITPVTDYVSGS-
                                                                                                                                                                                                                                                                                                                                                246;
                                                                                                                                                                                                                                                                                      4.2%; Score 171; DB 2; Length 908; 20.7%; Pred. No. 0.021; ive 98; Mismatches 272; Indels 2
Reference number: AB1807; MUID:21595285; PMID:11759840
Accession: AE2254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---GATWIRIWDWISY-----PNRSL---
                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 20.7°
Matches 161; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                    ::|:| | :
LSAEVIQQNNI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 1 1
                                                                                                                                                                                                                                                                                                                                                                                                    71
                                                                                                                                                                                                                                                                                                                                                                                                                                                            131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B8547
A;Status: preliminary
A;Rolecule type: DNA
A;Residues: 1-5188 <STO>
A;Cross-references: GB:AE005174; NID:g12513368; PIDN:AAG54838.1; GSPDB:GN00145; UWGP:Z061
C;Genetics: 4605 4109 4314 VIMVGALNWTYSDSGLVSGAYTYSARVVDLAGNITSSSDF----VLTVDTSIPTTLAQITS 4370 4371 QTTRDTTPIISGVITAALASGQYVEVVINGKTYTSEPGGAVVVDPAHNTWYVQLPDTDAL 4430 :| :| | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 4655 4656 VQHTYNLNNYYTLSSLINQGNGTFVWGQNTTNTFLSGAGSGAMSSSVSMTWADFDGDGDM 4715 4716 DLFLPASOGRANYGSLLFNTNGVLGCPVAVGATATTYASQFSLAVDWNHDGLMDIARIAQ 4775 47; 4431 TVSATAYTVTAQVKSSAGNGNNANISNGTVTVNAAIDYTPTWTTASKTTAW---GLTYGL 4487 415 176 522 --- OISWW---PDT--- 325 GQIHIAPMVKGLEETAVNDLISPPSGAPLIS----ALGDLGGFT-HADVTAVPSTIFTSP 470 86 27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG ------DNILYF-----GAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN 4160 ISTDTGSSATDFITSDTMLTLTGSLGAGLASGEVAQISLDSGATWTTLTTNGTQWTYTDS 177 PTDTTG---YQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPN---NPVFWSRDGGATW ---IJFRSTDGGATWTRIWDWTSYPNRSLRYVL DISAEPWLTFGVOPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG ----WSYVDGRTLTNGT--TTWQVRVV--DLAGNV-GATSSQSALIDTVNPAQVLTIAS : | | | :: | | | 4567 IQVIVGTLITWYGSI------VAFDKEGDGYLDFWIGDAGGPDSNT V----NPKTFYALS-----NGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGK-EG 4606 FLWNNAGTLVGNSTTSNSGGSATVGG------AVTGYLSLNEGSGVDLNNDGRIDL -----WKFS---VTSGTWT----RISPVPSTDTANDYFGYSGLTIDRQHPNTIMVAT-471 V-FTTGTSVDYAELNPSIIVRAGSFDPSSOPNDRHVAFST--DGGKNWFQGSEPG---------GVTTGGTVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDR 401; Length 5188; 1083 GANEFAQISTDNGA------TWVN-----VTVAADSLN-MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNN-QAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDG-----SSGDV--Indels DB 2; 4.2%; Score 170.5; Di 20.3%; Pred. No. 0.22; ive 82; Mismatches 3 DLWLAASSG-----LYHSTNG----Conservative 626

Page 9

6 6 6 6

y 642 -GSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG 690	Ov 471 V-FTTGTSVDYAELNPSIIVRAGSFDPSSOPNDRHVAFSTDGGKNWFOGSEPG 522
: : :	
Y 691 TTWVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGD 733 ::: : : b 4828 SVFLSRNTNTVSYGTSLHLRITDPNGINVYYGN 4860	OY 523GVTTGGTVAASADGSREVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDR 576 Db 4709 FLWNNAGTLVGNSTTSNSGGSATVGGAVTGYLSLNEGSGVDLNNDGRIDL 4758
ESULT 14 . 90696 ypothetical protein EC80542 [imported] - Escherichia coli (strain 0157:H7, substrain RI	Qy 577 VNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGK-EG 625
Escherichia coli Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001 11: F96696	Qy 626 DIWLAASSGLYHSTNG
; hayashi, 1:; makino, N.; Unnishi, M.; Kurokawa, K.; ishii, K.; Yokoyama, K.; Han, C.G. Baswara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. NA Res. 8, 11-22, 2001 Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno	Qy 642 -GSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG 690
Reference number: A99629; MUD:21156231; PMID:11258796 .Status: preliminary .Wolecult type: DNA .Residinar 1-5791 Alax	Qy 691 TIWULINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGD 733 : : :
CONTROLL OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF TH	RESULT 15 A28461 Profesory fore profess premy sor cartilage - rat
Cuery Match Query Match 4.2%; Score 168.5; DB 2; Length 5291; Best Local Similarity 20.1%; Pred. No. 0.3; Matches 190. Conservative 82. Wematches 270. Tabals 401. Gars 46.	N.Alternate names: aggrecan C.Species: Rattus norvegicus (Norway rat) C.Species: Jun-1989 #sequence revision 30-Jun-1989 #text_change 21-Jan-2000
27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGMNNWGYNGVSIAADPINTNKWWAAVG BILL STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER STANDER ST	R; Doege, K.; Saeaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y. J. Biol. Chem. 262, 17757-17767, 1987 A; Tile: Complete primary structure of the rat carliage proteoglycan core protein deduce A;Reference number: A92623; MUD:88087070; PMID:3693370
87 MYTNSWDPNDGAILRSSDQGATWQITPLPPKLGGNWPGRGMGERLAVDPNN 137	
	J. Biol. Chem. 263, 10040a, 1988 A;Reference number: A30069 A;Contents: annotation; revision to residue 698 R;Doege, K.; Fernandez, P.; Hassell, J.R.; Sasaki, M.; Yamada, Y.
	J. Biol. Chem. 261, 8108-8111, 1986 A;Title: Partial CDNA sequence encoding a globular domain at the C terminus of the rat of A;Reference number: A23835; MUID:86250698; PMID:2424893 A;Accession: A23835
	A;Molecule type: mRNA A;Residues: 1856-2124 <do2> A;Cross-references: GB:M13518; NID:g206104; PIDN:AAA41836.1; PID:g206105 R:Noame, P.J.: Christner, J.E.: Baker, J.R.</do2>
WKESVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVAT	J. Biol. Chem. 262, 17768-17778, 1987 A;Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-termi A;Reference number: A28453; MUID:88087071; PMID:3693371
OTTERPREDET SCRITTER AND AND AND AND AND AND AND AND AND AND	A,Molecula type: protein A,Residues: 20-37,'W',39-60,'E',62-64,'X',66-69;70-83;84,89-148,'L',150-238,'S',240,'A',; C,Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EGE
TVSATAYTVTAQVKSSAGNGNNANISNGTVTVYVAAIDYTPWTTASKTTAW	F;1.19/Domain: signal sequence #status predicted <sig>F;1.19/Domain: signal sequence #status predicted <mat>F;20-2124/Product: proteoglycan core protein #status predicted <mat>F;4-135/Domain: immunoglobulin homology <pre>cIMM>F;170-247/Domain: link protein repeat homology <pre>cINK1></pre></pre></mat></mat></sig>
	F;268-349/Domain: link protein repeat homology <lnk2> F;504-581/Domain: link protein repeat homology <lnk3> F;602-683/Domain: link protein repeat homology <lnk4> F;1914-2034/Domain: C-type lectin homology <lch></lch></lnk4></lnk3></lnk2>
y 416 GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFT-HADVTAVPSTIFTSP 470 	F;2041-2097/Domain: complement factor H repeat homology «FHD» F;126,239,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status predic Query Match 4.2%; Score 168; DB 2; Length 2124;

8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6

```
| | | | :: | | : | | | 1426 SGPSGFPVLSGTTPGTPDLASGAMSGSGDSSGITFVDTSLIEVTPTTFREEEGLGSVEL 1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44;
                                                                          1038 SGIEDITVLPTGRENLETSASGVEDVSGLPSGKEGLETSASGI--EDISVFPTEAEGLET 1095
                                                                                                                                                                                                                   1096 SASGGYVSGIPSGEDGTETSTSGVEGVS---GLPSGGEGLETSASGVEDLGLPTRDSLET 1152
                                                                                                                                                                                                                                                                                                                                                            1211 VGSASGALDFCKLPSGTLGSGQTPEASGLPS-GFSGEYSGVDIGSGPSSGLPDFSGLPSG 1269
                                                                                                                                                                                                                                                                                                                                                                                                                           1330 SGQTSGS-----LDVSGETSGFPDVSGQPFGSSGTGEGTSGIPEVSGQAVRSPDTTEI 1382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | :| | :| | :: : : 1486 SGLPSGETDLSGTSGWTDVSGQSSGAIDSSGLISPTPEF-SGLPSGVEVSGEVSGVETG 1544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        527 GGTVAASADGSRFVWAPG-----DPGQPVVYAVGFGNSWAAS------QGVP 567
                                                   66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPPKLGGNMPGR 125
                                                                                                                       126 GMGERLAVDPNNDNIL-----YFGAPSGKGLWRSTDSGATWSQMTNFPD----- 169
                                                                                                                                                                                          170 --VGTYIAN-PIDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDG 226
                                                                                                                                                                                                                                                                                                                                -GSSGDVWKF-----YSGTWTRISPVPSTDTANDYFG------YSGL--- 303
                                                                                                                                                                                                                                                                                                                                                                                                     ------QISWWPDTIIFRSTD--GGATWTRI 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----PVPSPKLGWM 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                488 -IVRAGSFD------PSSQPNDRHVAFS-----TDGGKN-----WFQGSEPGGVTT 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----NPKTFYALSNGTFYRSTDGGVTFQP---VA 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          605 AGLPSSGA------VGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV-N 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             657 VGFGKSAPG---SSYPAVFVVGT-----IGGVTGAYRSDDCGTTWV----LINDDQ 700
                                                                                                                                                                                                                                                              GATWQAVPGAPTG-----FIPHKGVFDPVNHVLY----IATSNTG------GPYD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GDLGGFTHADVTAV---PST------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         382 DEAMAIDPFNSD-----RMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                466 -----IFTSPVFTTGTSVDYAELNPSI----
                     Gaps
Best Local Similarity 20.5%; Pred. No. 0.096;
Matches 190; Conservative 102; Mismatches 311; Indels 322;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     341 WDWTSYPNRSLRYVLDISAEPWLTFGVQPNP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1828 TEISGEW----SDHTSEVNVTVST 1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 701 HQY-GNWGQAITGDHANLRRVYIGT 724
                                                                                                                                                                                                                                                                                                                                                                                                    --TIDROHPNTIMVAT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        568 ANAQIRSDRV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             436 ISPPSGAPLISAL--
                                                                                                                                                                                                                                                              227
                                                                                                                                                                                                                                                                                                                                                                                                    304
                                                                              d
                                                                                                                       ò
                                                                                                                                                     g
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                            쉽
                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               유
                                                     8
                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

Search completed: May 11, 2004, 12:10:24 Job time : 40.7001 secs us-09-917-376-3.rsp

```
5.1.6
Compugen Ltd.
GenCore version (c) 1993 - 2004
           Copyright
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 11:41:35 ; Search time 20.1745 Seconds (without alignments) 1909.933 Million cell updates/sec

US-09-917-376-3

4036 1 ATTQPYTWSNVAIGGGGFVD......XIGTNGRGIVYGDIGGAPSG 740 Perfect score: Seguence:

Gapop 10.0 , Gapext 0.5 **BLOSUM62** Scoring table:

141681 segs, 52070155 residues

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

SwissProt_42:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	SUMMARIES	Description	P07897 rattus norv	candida	P49791 rattus I	QUIA_ACICA Q59086 acinetobact				Q61282	P10342	P26501	P25927		09xd78			QHED_COMTE Q46444 comamonas t	P49793	P48839	Q02834	005542	P75780	P26225		099715		YY46_ANASP Q8yril anabaena sp		Q9eprs 1	P50401	BPAR1 Q9g0b5 bacteriopha		HIMAN P37198 homo sapien
		DB	: -	-	П	Н	~	-	٦	٦	٦	Н	Н	Н	Н	н	Н	-	Н	Н	Н	Н	н	Н	н	-4		Н	Н	Н	Н	Н	-	_
1		Length	2124	1119	1468	809	5703	1150	1196	2132	176	176	1953	2109	190	342	747	708	937	995	647	757	760	1045	2333	3063	955	1526	827	1159	872	1103	4	Ľ
1	d	Query Match	4.2	4.2	4.1	3.9	3.8	3.8	3.8	3.7	3.7	3.7	3.7	3.6	3.6	3.6	3.6	3.6	3.6	3.5	3.5					נייו	(*)	۳,			,		1.,	٣
A COURTY MARCH Length March Length 4.2 1119 4.1 1119 11.0 11.0 11.0 11.0 11.0 11.0 11		Score	168	167.5	166	157	154.5	153	152	149	148.5	148.5	148.5	146	145.5	145	145	144.5	143.5	142.5	142	142	142	139.5	139.5	139	138.5	138.5	138	138	137.5	137.5	137.5	136.5
Ouery Match Length 4.2 2124 4.2 2124 4.2 2124 4.2 3.8 3.8 3.8 3.7 3.9 3.7 3.6 3.7 3.6 3.7 3.6 3.6 3.6 3.6 3.6 3.6 3.6 3.6 3.6 3.6		Result No.		7	m	4	S	9	7	8	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32

P52948 homo sapien	Q51845 porphyromon	Q97z97 sulfolobus	P27410 rabbit hemo	P08696 clostridium	P29767 clostridium	P02674 petromyzon	Q07284 epstein-bar	P16397 bacillus su	Q8ixj9 homo sapien	_	P52591 rattus norv	
NU98 HUMAN	HGA2_PORGI	YA33 SULSO	POLN RHDV	BCN5_CLOPE	NANH CLOSE	FIB1 PETMA	VGP3 EBVA8	SUBF_BACSU	ASX1_HUMAN	VG37 BPT4	P121_RAT	
-	Н	٦	Н	Н	Н	Н	Н	ч	ч	٦	-	
1729	2628	725	2344	890	1014	996	886	1433	1541	1026	1199	
3.4	3.4	3.4	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	
136.5	136	135.5	134.5	134	134	133.5	132	132	132	131.5	131.5	
34	35	36	37	38	39	40	41	42	43	44	45	

ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -i- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate chains, N-linked and O-linked oligosaccharides.
-i- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-i- SIMILARITY: Contains 4 like domains.
-i- SIMILARITY: Contains 1 C-type lectin family domain.
-i- SIMILARITY: Contains 1 Sushi (SCR) domain.
-i- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN: Two globular domains, G1 and G2, comprise the amino terminus of the proteoglycan, while another globular region, G3, makes up the COOH terminus. G1 contains link domains and thus consists of three disulfide-bonded loop structures designated as the A, B, B motifs. G2 is similar to G1. The keratan sulfate (KS) and the chondroitin sulfate (CS) attachment domains lie between G2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=86250698; PubMed=2424893;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
Doege K., Fernandez P., Hassell J.R., Sasaki M., Yamada Y.;
J. Biol. Chem. 261:8108-8111(1986).
J. Biol. Chem. 261:8108-8111(1986).
I. FUNCTION: This proteoglycan is a major component of extracellular matrix of cartilagenous tissues. A major function of this protein is to resist compression in cartilage. It binds avidly to hyluronic acid via an amino-terminal globular region. May play a regulatory role in the matrix assembly of the cartilage.
I. SUBUNIT: Interacts with FBLNI (By similarity).
-: SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
                                                                                                       01-FEB-1991 (Rel. 17, Last sequence update)
15-WAR-2004 (Rel. 43, Last annotation update)
Aggrecan core protein precursor (Cartilage-specific proteoglycan core
protein) (CSPCP).
                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Rattus.
                                                                                                                                                                                                                                                                                                                                                        MEDLINE=88087070; PubMed=3693370;
Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
Mcomplete primary structure of the rat cartilage proteoglycan protein deduced from cDNA clones.";
J. Biol. Chem. 262:17757-17767(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Doege K., Sasaki M., Horigan E., Hassell J.R., Yamada Y.;
J. Biol. Chem. 263:10040-10040(1988).
                                            2124 AA
                                            PRT;
                                                                                       01-AUG-1988 (Rel. 08, Created)
                                          STANDARD;
                                                                                                                                                                                                                           Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                      NCBI_TaxiD=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REVISION TO 698.
                                                                                                                                                                                                        AGC1 OR AGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and G3
                                          PGCA RAT
RESULT 1
                       PGCA_RAT
```

Thu May

```
CARBOHYD
CARBOHYD
                                                                                                             Local
    FTF
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                 g
                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                    ठे
                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . .) (POTENTIAL)
. .) (POTENTIAL)
. .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00032; CCP; 1.

SMART; SM00034; CLECT; 1.

SMART; SM00406; IGV; 1.

SMART; SM00406; IGV; 1.

SMART; SM00406; IGV; 1.

PROSITE; PS00615; C TYPE IECTIN 1; 1.

PROSITE; PS00209; IG INE; 1.

PROSITE; PS00209; IG MHC; 1.

PROSITE; PS01241; INIK; 4.

Glycoprotein; Protecoglycan; Lectin; Signal; Sushi; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
AGGRECAN CORE PROTEIN.
IG-LIKE V-TYPE.
LINK 1.
LINK 2.
LINK 3.
LINK 3.
LINK 4.
C-TYPE LECTIN.
SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G1-A.
G1-B.
G1-B.
G2-B.
G2-B.
CS-1.
CS-1.
CS-1.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
G3-S.
                                                                                                                                                                Interpro; IPR00233; Antifreezell.
Interpro; IPR00233; Ig-like.
Interpro; IPR003006; Ig-MfC.
Interpro; IPR003506; Ig-V.
Interpro; IPR001304; Lectin_C.
Interpro; IPR00334; SGXXSG.
Interpro; IPR00334; SGXXSG.
                                                                                                       EMBL; M13518; AAA41836.1; -.
EMBL; J03485; AAA21000.1; ALT_SEQ.
PIR; A32623; A28452.
HSSP; P98066; 1TSG.
                                                                                                                                                                                                                                                                                     Pfam; PF00047; ig; 1.
Pfam; PF00059; lectin c; 1.
Pfam; PF00039; SGXXSG; 55.
Pfam; PF00199; SIGNI; 1.
Pfam; PF00193; Xlink; 4.
PRINTS; PR00156; ANTFREZEII.
PRINTS; PR01265; LINKMODULE.
ProDom; PD000918; Link; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          581
683
2036
2098
140
247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   682
798
11226
11909
2124
133
246
220
348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1925
2034
2026
2084
2097
126
333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          504
602
1910
2040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
```

```
096 SASGGYVSGIPSGEDGTETSTSGVEGVS---GLPSGGEGLETSASGVEDLGLPTRDSLET 1152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ||: : || : : || : || 111 VGSASGALDFGKLPSGTLGSGTPEASGLPS-GFSGEYSGYDIGSGPSSGLPDFSGLPSG 1269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGQPSGFPVLSGTTPGTPDLASGAMSGSGDSSGITFVDTSLIEVTPTTFREEEGLGSVEL 1485
                                                                                                                                                                                                                                                                       1330 SGQTSGS-----LDVSGETSGFFDVSGQPFGSSGTGEGTSGIPEVSGQAVRSPDTTEI 1382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1383 SELSGLSSGQPDVSGEGSGILFGSGQSSGITS------VSG-ETSGISDL 1425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EVTLITSELVEGVTEPTVSQELGHGPSMTYTPRLFEASGEASASGDLGGPVTIFPGSGVE 1722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1486 SGLPSGETDLSGTSGMVDVSGQSSGAIDSSGLİSPTPEF-SĞLPSGVAEVSGEVSGVETG 1544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SSLSSGAFDGSGLVSGFPTVSLVDRTLVESITLAPTAQEAGEGPSSILEFSGAHSGTPDI 1604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TDLEMTTPGTEVSSNPWTFQEGTREGSAAPEVSGESSTTSDIDAGTSGVPFATPMTSGDR 1827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----QISWWPDTIIFRSTD--GGATWTRI 340
                                                                                                                                                                                                                                          NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GSSGDVWKF-----SVTSGTWTRISPVPSTDTANDYFG------YSGL--- 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----NPKTFYALSNGTFYRSTDGGVTFQP---VA 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGLPSSGA-----VGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV-N 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VGFGKSAPG---SSYPAVFVVGT-----IGGVTGAYRSDDCGTTWV----LINDDQ 700
                                                                                                                                                                                                                                                                                                                                                                                                                                 --VGTYIAN-PIDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDG
                                                                                                                                                                                                                                                                                                                                     -----YFGAPSGKGLWRSTDSGATWSQMTNFPD-----
                                                                                                                                                                                                                                                                                                                                                                    | : | | : | | : | | SGIEDITVLPTGRENLETSASGVEDVSGLPSGKEGLETSASGI--EDISVPPTEAEGLET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GATWQAVPGAPTG-----FIPHKGVFDPVNHVLY----IATSNTG------GPYD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---- PVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEAMAIDPENSD-----RMLYGTGATLYATNDLTKWDSGGOIHIAPMVKGLEETAVNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       436 ISPPSGAPLISAL-----GDLGGFTHADVTAV---PST------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              488 -IVRAGSFD------PSSQPNDRHVAFS-----TDGGKN-----WFQGSEPGGVTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGTVAASADGSRFVWAPG-----DPGQPVVYAVGFGNSWAAS------QGVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----IFTSPVFTTGTSVDYAELNPSI-----
387 N-LINKED (GLCNAC. ..) (POTENTIAL)
611 N-LINKED (GLCNAC. ..) (POTENTIAL)
667 N-LINKED (GLCNAC. ..) (POTENTIAL)
1842 N-LINKED (GLCNAC. ..) (POTENTIAL)
7, 221117 MW; E30BBE61593A34B1 CRC64;
                                                                                                                                                                                            Matches 190; Conservative 102; Mismatches 311; Indels 322;
                                                                                                                                             4.2%; Score 168; DB 1; Length 2124; 20.5%; Pred. No. 0.06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WDWTSYPNRSLRYVLDISAEPWLTFGVQPNP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1828 TEISGEW-----SDHTSEVNVTVST 1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HQY-GNWGQAITGDHANLRRVYIGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASVPEGSSDPSAYPEAGVGVSAAP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --TIDRQHPNTIMVAT-----
                                                                                                                                                                                                                                                                                                                                       GMGERLAVDPNNDNIL
                                                                                                 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        568 ANAQIRSDRV-
                                                                                                                                                                      Similarity
  387
611
667
1842
2124
                                              CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                   1038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1663
                                                                                                                                                                                                                                             99
                                                                                                                                                                                                                                                                                         979
                                                                                                                                                                                                                                                                                                                                     126
                                                                                                                                                                                                                                                                                                                                                                                                                                 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            701
                                                                                                                                                Query Match
```

us-09-917-376-3.rsp

```
1468 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             970 NSEPTTŠTAASTSTDIENETIATTGSV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                      WKFSVTSGTWTRISPVPSTD-
          Best Local Similarity 18.7
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1067 SOYNSDTOOTT 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        701 HOYGNWGOAIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
N153_RAT
ID N153_RAT
                                                                                                                                                                                                                                                                                     273
                                                                                                                                                                                                                                                                                                                            307
Query Match
Best Local
                                                          셤
                                                                               ઠ
                                                                                                                                                               ઠ
                                                                                                                                                                                                    8
                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                유
                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                             gene
                                                                                                       Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
NCBI_TaxID=5476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                              STRAIN=1161;
MEDILINE=98309840; PubMed=9644209;
HOVER L.L., Payne T.L., Bell M., Myers A.M., Scherer S.;
"Candida albicans ALS3 and insights into the nature of the ALS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                              Curr. Genet. 33:451-459(1998).
-!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC. . .) (POT MW; 6A3FB3FC8C879A71 CRC64;
                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
AGGLUTININ-LIKE PROTEIN 3.
                                                                                                                                                                                                                                                                                                                                                                                                  10 X 36 AA TANDEM REPEATS
                                                                                                                                                                                                                                  -!- PTM: N-glycosylated and O-glycosylated (Potential)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                          (Rel. 38, Last sequence update) (Rel. 38, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                    Repeat; Signal.
                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                              Agglutinin-like protein 3 precursor.
                                                                                                                                                                                                                                                                                                                                              Interpro; IPR008440; Candida ALS.
Pfam; PF05792; Candida ALS; 1.
Cell adhesion; Glycoprotein; Repe
                                                15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last seq
15-JUL-1999 (Rel. 38, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119927
                                                                                                                                                                                                                                                                                                                                      EMBL; U87956; AAC39486.1;
                                                                                                  Candida albicans (Yeast)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ä.
                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1050
1061
1119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     450
                            ALS3 CANAL
074623;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
```

```
-----EASSPIISSSADETTTV 1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1014 TTTAESTSVIEQPTNNNGGGKAPSATSSPSTTTTANNDSVI-----TGTTSTNQSQSQ 1066
                                                            42;
                                                                                                                                                                            313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               807 -SYTTATIVTAPPGGIDIVIIYDIMSSSEISSFSRPHYTNHTILMSITWVIETKTIITETS 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               866 CEGDKGCSWVSVSTRIVTIPNNIETPMVTN-----TVDSTTTESTSQSPSGIFSESGVS 919
                                                                                                                                                                                                                                                                                                                                                         116 -----FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGAT---- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 806
                                                                                                                                                                                                                                                                                               363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --TTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         183 YOSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPV----FWSRD--GGATWQAVPG 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            236 APTGFIPHKGVFDPVNHVL-----YIATSNTGGPYDG------SSGDV 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----1D 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        397 ------WDSGGQIHIAPMVKG 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    920 VETESSTVTTAQTNPSVPTTESEVVF-----TTKGNNENGPYESPSTNVKSSMD-E 969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWS 646
                                                                                                                                                                                                                                                                       121 IDSIDTVIVQVPSPNPTVTTTEYWSQ--SFATTTTITGPPGNTDTVLIREPPNHTVTTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RQHPNTIMVATQISWW---------PDTIIFRSTDGGATWTRIWDWTSYPNRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         700 FATTITVTAPPGGIDIVIIREPPNHTVITIEY-WSQSYATITIITAPPGEIDTVLIREPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 TSVD----YAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           332 ASADGSRFVWAPGDPGQP----VVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              647 AITGVSSAV-----NVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQ
                                                                                                                   7 TWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGG-----MYRWDAANGRW--IPL-LDWV
                                                                                                                                                                         264 TYKNVPAGYRPFVDAYI-----SATDVNSYTLSYANEYTCAGGYWQRAPFTLRWT
                                                                                                                                                                                                                                   59 GWNN--WGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQI-TPLP
                                                                                                                                                                                                                                                                                                                                                                                                               364 TTTITTSYVGVTTSYSTKTAPIGETATVIV---DIPYHTTTTVTSKWTGTITSTTTHTNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     351 LRYVLDISAEPWLTFGV----QPNPPVPSPKLGWMDE-----AMAIDPFNSDRML----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              759 NHTVTTTEYWSQSYATTTTIIAPPGETDTVLIREPPNPTVTTTEYWSQ-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           127 LEETAVNDLISPPSGAPLI----SALGDLGGFT-----HADVTAVPSTIFTSPVFTTG
                                                         303; Gaps
Length 1119;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------TANDYFGYSGLT-
                                                         98; Mismatches 340;
DB 1;
Score 167.5; Di
Pred. No. 0.03;
```

```
-----SVTPFTYKTTEAKKEDASATKGGFT-FGKVDSAALSSPSMFVLGRTEEKQ 1100
                                                                                                                                                                                                                                                                                                                                                                  1137 -ESSSKPTFSFSVAK----PSVK-----ESDQLAKATFAFGNQTNTTT--DQ 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1283 TFGATTTSSSSGSFFVFGTGHSAPSASPAFGANOTPTFGOSOGASO------PN 1330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1331 PPSFGSISSSTALFSA-GSQPVPPPTFGTVSSSSQPPVFGQQPSQS-----AFGSGTANA 1384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SSVFQFGSSTTNFNFTNNNPSGVFTFGASPSTPAAAAQPSGSGGFSFSQSPASFTVGSNG 1444
                                                     997 TLGQQEKKEELPQSSSAGFSFGAGVANPSSAAI---DTTVTSENKSGFNFGTIDTKSV-- 1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GA---AKPAFSFLNSS-----SSSSSTPATSSSASIFGSSTSSSSPPVAAFVFGQASNP 1227
                                                                                                                                                                                                                                                                                                                                                                                                                        GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDL-GGFTHADVTAVPSTIF---TSP 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TTGGTVAASADGSREVWAPGDPGQPVVYAVG-----FGNSWAASQGVPANAQIRSDRVN 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    579 PKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHS 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       639 TNGGSSWSAITGV-----SSAVNVGFGKSAPG------SSYPAVFVVGTIG 678
     192 WVAFDKSSSSLGQASKTIF---VGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFD 248
                                                                                                                                                                                                            --GGATWTRIWDWTSYPNRSLRYV 354
                                                                                                                                                                                                                                                                                                                  355 LDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VFTT--GISVDYAELNPSIIVRAGSPDPSSQPNDRHVA----FSTDGGKNWFQGSEPGGV 524
                                                                                                        249 PVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDR-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Unusual ancestry of dehydratases associated with quinate catabolism in Acinetobacter catcoaceticus.";
J. Bacteriol. 177:5978(1995)
J. Bacteriol. 177:5978(1995)
J. Bacteriol. 277:5978(1995)
J. CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-dehydroquinate + reduced pyrroloquinoline-quinone.
J. CATALYTIC ACTIVITY: Shikimate + pyrroloquinoline-quinone = 3-dehydroshikimate + reduced pyrroloquinoline-quinone.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of Acinetobacter calcoaceticus
for quinate-shikimate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Moraxellaceae; Acinetobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Quinate/shikimate dehydrogenase [Pyrroloquinoline-quinone]
(EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
                                                                                                                                                                                                                                                               1101 ÓEPVT-----STSLVFGKKADNEEPKCOPVFSFGNSEOTKD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND CHARACTERIZATION STRAIN=BD413 / ADP1;
                                                                                                                                                                                                              OHPNTIMVATOISWWPDTIIF-RSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Elsemore D.A., Ornston L.N.;
"The pca-pob supraoperonic cluster contains quiA, the structural gene dehydrogenase.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=95095936; PubMed=8002591;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=96011389; PubMed=7592351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteriol. 176:7659-7666(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Acinetobacter calcoaceticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Elsemore D.A., Ornston L.N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-20 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=BD413 / ADP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=471;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QUIA ACICA
                                                                                                                                                                                                              308
                                                                                                                                                           1052
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1385
                                                                                                                                                                                                                                                                                                                                                                                                                        415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              471
                                                                                                        ð
                                                                                                                                                   g
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        164 MINF-------PD----PD-----PD------VGTYIANPIDITGYQSDIQGVV 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GATWQITPLP--FKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQ 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  259; Indels 186; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FUNCTION: Possible DNA-binding subunit of the nuclear pore complex (NPC). The repeat-containing domain may be involved in anchoring components of the pore complex to the pore membrane. SUBUNIT: Interacts with SENP2 (By similarity).
SUBCELLULAR LOCATION: Nuclear pore complex. Located to the terminal ring structure of the nucleoplasmic cage.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEATS
                          01-OCT-1996 (Rel. 34, Created)
1-OCT-1996 (Rel. 34, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Nuclear pore complex protein Nup153 (Nucleoporin Nup153) (153 kDa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARIȚY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATY
SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of the
                                                                                                                                                                                                                                                                                                                                                                                                               "A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm."; Cell 72:29-38(1993).
                                                                                                                                                                                                                                                                                                               AND SEQUENCE OF 602-613; 622-645 AND 971-993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PSO1358; ZF_RANBP2_1; 4.
PROSITE; PSS0199; ZF_RANBP2_2; 4.
Nuclear protein; Transport; Repeat; Zinc-finger; DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 166; DB 1; Length 1468;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN: Contains F-X-F-G repeats.

PTM: Phosphorylated in interphase, hyperphosphorylated mitosis. May play a role in the reversible disassembly nuclear pore complex during mitosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MW; C3DFD9697C556A7C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUP1, NSP1, POM 121 AND MAMMALIAN P62.
SIMILARITY: Contains 4 RANBP2-type zinc fingers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21.8%; Pred. No. 0.051; ive 71; Mismatches 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLY-RICH.
POLY-GLY.
RANBP2-TYPE 1.
RANBP2-TYPE 2.
RANBP2-TYPE 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A44345; A44345.
InterPro; IPR004325; Nucleoporin_FG.
InterPro; IPR01876; Znf RanGDP.
Pfam; PF03093; Nucleoporin_FG.
Pfam; PF00641; Zf-RanBP; 4.
SMART; SM00547; ZnF RBZ; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; LO6821; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND SEQUENCE (STRAIN-Buffalo, TISSUE-Liver; MEDLINE-93137325; Pubmed-8422679;
                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata;
Mammalia; Butheria; Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 846 875
1468 AA; 152823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 21.8
Matches 144; Conservative
                                                                                                                                                                                   Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                          Sukegawa J., Blobel G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15
446
687
750
818
875
                                                                                                                                                                                                                                                            NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Phosphorylation
                                                                                                                                nucleoporin)
NUP153.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZN_FING
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
ZN_FING
ZN_FING
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
STATE THE MEN AND DEPTH OF THE PROPERTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF THE PARTY OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
```

us-09-917-376-3.rsp

```
619 MFVNDMRLGLWIQLIKQTPED--IKIQANGGEKVNTGM--GAVPMKGT-PYKVNKRRFMS 673
                          MEDLINE=99023932; PubMed=9804771;
Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
"Genomic organization of the human mucin gene MUCSB: cDNA and genomic sequences upstream of the large central exon.";
J. Biol. Chem. 273:30157-30164(1998).
THADVTAVP-----STIFTSPVFTTGTSVDYA-----ELNPSIIVRAGSFDPSSQ-
                                                                                                                                                                                                       APGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF-
                                                                                                                                                                                                                                                          674 ALGIP------COKPPFGTMTAIDMKTRQVAWQVPLGTIQDTGPMGIKMG
                                                                                                                                                                                                                                                                                                             --QPVAAGLPSSGAV----GVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGV-S
                                                                                                                                                                                                                                                                                                                                                                718 LKAPI - GMPTIGGPMATQGGLVFFAA - - - TQDYYLRA - - - - FNSSNGKELWKARLPVGS
                                                                                                        --------PNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternate subdomains resulting in a super-repeat. Structural evidence for a 11p15.5 gene family.";
J. Biol. Chem. 272:3168-3178 (1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIINE=99009274; PubMed=9790959;
Gfiner G.D., Nunes D.P., Keates A.C., Afdhal N.H., Troxler R.F.;
"The amino-terminal sequence of MUC58 contains conserved
multifunctional D domains: implications for tissue-specific mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Craniata, Vertebrata; Buteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9HCB4; Q00447; Q00573; Q14985; Q15494; Q95291; Q95451; Q14881; Q09552; Q9UE28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=97166151; PubMed=9013550; Desseyn J.-L., Guyonnet-Duperat V., Porchet N., Aubert J.-P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W.738-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (Hig
MuCin 5B precursor (Mucin 5 subtype B, tracheobronchial) (Mig
MUCSB OR MUCS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Salivary gland;
MEDINE=97292540; Pubmed=9147051;
Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
Hannibal J., Clausen H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chen Y., Di Y.P., Wu R.; "Molecular cloning of the amino-terminal and 5'-flanking human MUGSB mucin gene."; Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                      SAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                      169 QGTPMSYMSPKTGKQY----VVVSAG---GARQSPDHG 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Siochem. Biophys. Res. Commun. 251:350-355(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1326-4895 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 4057-4480 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bukaryota, Metazoa; Chordata;
Mammalia; Butheria; Primates;
MCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 40-1324 FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 1-1594 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-1325 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             rissum=Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unctions.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Laine A.;
  455
                                                                                                        200
                                                                                                                                                                                                               542
                                                                                                                                                                                                                                                                                                                     601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MUSB HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HID DITION OF SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE 
                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                   셤
                                                                                                        ઠ
                                                                                                                                                 유
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                          음
                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                302 GLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATW-----TRIWDWTSYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       131 GGLTVLGMLGGLYGMFIPHETVKASGEELPLVPVDPAKKQVNWDHYGNDAGGSRFVALDQ 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   76 INTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                296 TECPRRVYTNTVD-GRLIAVNADTGARCKDFGVNGTV----NLHEGLGENTKAPRFEVT 349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         250 VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDT------ANDYFGYS 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30 GILYVRTDIGGMY------RWDAANGRWIPLL-----DWVGWNNWGYN--GVVSIAADP 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          161 ----WSQMTNFPDVGTYIANPTDTTGYQSD--IQGVVWVAFDKSSSSLGQASK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      208 -----TIFVG-----VADPNNPVFWSRD---GGATWQAVPGAPTGFIPHKGVFDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      350 SAPTIAGTTIVVGSRIADNVAADMPGGVIRAYDVITGKLRWAFDPRNP-----DP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          400 -NYVL-------KPGEIYKRSSTN-SWAAMSYDPOMNTVFLPMGSSSVDVWGGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   445 RTAADHKY-NTSVLALD------ATTGKEKWVYNTVHNDLWDF------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISAEPWLT-FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    415 GGOIHI-----APMVKGLEE-----TAVNDLISPPSGAPLISALGDLGGF
                          PATHWAY: Quinic acid catabolic pathway; first step. This pathway allows growth of bacteria with quinate by its conversion to protocatechuate and subsequent metabolism by the beta-ketoadipate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.9%; Score 157; DB 1; Length 809;
20.5%; Pred. No. 0.081;
(ve 90; Mismatches 254; Indels 306; Gaps
                                                                                                                                                            -!- INDUCTION: By protocatechuate.
                                                                                                                                 SUBCELLULAR LOCATION: Integral membrane protein (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00564; PQQ; 3.
PROSITE; PS00363; BACTERIAL PQQ_1; FALSE_NEG.
PROSITE; PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      71F67CEBEA62BFCB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A55547; A55547.
InterPro; IPR001479; Bac_PQQ.
InterPro; IPR002372; Bac_PQQ_repeat.
Pfam; PP01011; PQQ; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |: :| | :: | | : :|
481 DLPMOPSLVDFPMKDGTTKPAVVIG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            136 NNDNILYFGAPSGKGLWRS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88196 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L05770; AAC37161.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
88
110
147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      809 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
```

region of the

(High

```
930
1277
1293
11557
11557
11755
21757
21757
3948
4745
4901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                        DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                SULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                               SULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          DISULPID
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97347489; PubMed=9201995;

A Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;

Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;

Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;

"Genomic organization of the 3 region of the human MUC5B mucin.";

"J. Biol. Chem. 272:16873-16883 (1997).

-!- FUNCTION: Salivary mucin that is thought to contribute to the lubricating and viscoelastic properties of whole saliva.

-!- FUNCTION: Salivary mucin that is thought to contribute to the C.:- SUBCELLULAR LOCATION: Secreted.

-!- SUBCELLULAR LOCATION: Secreted.

-!- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also contributed and pancreas.

-!- FTM: Highly glycosylated.

-!- FTM: Highly glycosylated.

-!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.

-!- SIMILARITY: Contains 4 VWFC domains.

-!- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
"Identification of a major human high molecular weight salivary mucin (MG1) as tracheobronchial mucin MUC5B.";
                                                                                                                              Keates A.C., Numes D.P., Afdhal N.H., Troxler R.F., Offner G.D.; "Molecular cloning of a major human gall bladder mucin: complete C-terminal sequence and genomic organization of MUC5B."; Blochem. J. 324:295-303[1997].
                                                                                                                                                                                                                                               TISSUE=Sublingual gland;
MEDLINE=56125355; PubMed=8554565;
Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim P.G.;
"Molecular cloning of a novel high molecular weight mucin (MG1)
from human sublingual gland.";
Biochem. Biophys. Res. Commun. 217:1112-1119(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM: 600770; ...
GO; GO: 6005515; F: protein binding; IPI.
GO; GO: 6005515; F: protein binding; IPI.
InterPro; IPR005207; Cys. knot. C.
InterPro; IPR005904; PMP inhibitor.
InterPro; IPR002919; TIL_Cysrich.
InterPro; IPR001904; VWF C.
InterPro; IPR001007; VWF C.
InterPro; IPR001846; VWF D.
Pfam; PF00007; Cys. knot; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EWBL; AF107890; AAG33673.1; --
EWBL; AF086604; AAC67545.1; --
EWBL; AJ004862; CAA96167.1; --
EWBL; 272496; CAA96577.1; --
EWBL; X74955; CAA52910.1; --
EWBL; U3836; CAA51348.1; --
EWBL; U78552; AAC51344.1; JOINED.
EWBL; U78553; AAC51344.1; JOINED.
EWBL; U78551; AAC51343.1; --
EWBL; U95031; AA565151.1; --
EWBL; Y09788; CAA70926.1; --
GGDEW; HGNC;7516; MUC5B.
                                                                                                                MEDLINE=97293229; PubMed=9164870;
                                                                           4721-5703 FROM N.A.
bladder;
                                                                                                                                                                                                                             SEQUENCE OF 4809-5687 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 4859-5703 FROM N.A.
                                       3lycobiology 7:413-419(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00094; vwd; 4.
SMART; SM00214; VWC; 6.
SMART; SM00215; VWC_out; 4.
                                                                                                                                                                                                                                                                                                                                                                                                        rissum=Placenta;
                                                                           SEQUENCE OF 
FISSUE=Gall
                                                                           SEQUENCE
```

```
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /FTId=UAR 014123.

G -> E (IN REF. 2).
FPGLCN -> LDFCLCK (IN REF. 2).
E -> K (IN REF. 1).
R -> S (IN REF. 2).
E -> K (IN REF. 2).
E -> N (IN REF. 2).
E -> N (IN REF. 2).
E -> N (IN REF. 2).
E -> K (IN REF. 2).
E -> K (IN REF. 2).
E -> K (IN REF. 2).
E -> K (IN REF. 2).
E -> K (IN REF. 2).
C -> R (IN REF. 2).
C -> R (IN REF. 2).
C -> R (IN REF. 2).
C -> R (IN REF. 2).
C -> R (IN REF. 2).
C -> N (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 2).
C -> P (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (Caraman de Caraman de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                         TIL.
VWFD 2.
VWFD 3.
THR-RICH.
THR-RICH.
VWFC 2.
VWFC 3.
CTCK.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                         Polymorphism.
                                                                                                                           POTENTIAL.
MUCIN SB.
VWFD 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
SMART; SM00216; VWD; 4.

PROSITE; PS01185; CTCK_1; 1.

PROSITE; PS01225; CTCK_2; 1.

PROSITE; PS01208; VWFC_1; 2.

PROSITE; PS0184; VWFC_1; 2.

Glycoprotein; Repeat; Signal; PC SIGNAL 1 25 GN03

CHAIN 26 5703

MUC DOMAIN 329 386 TII

DOMAIN 426 580 VWI

DOMAIN 858 918 VWI

DOMAIN 1457 1603 THH

DOMAIN 1457 1603 THH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5178
5425
5528
5683
5646
5660
5676
5678
```

```
23;
                                                                                                                                                                                                                2956 SSTPGTTWILTEQTTAATTTATTGSTAIPSSTPGTAPPRVLTSQATTPT----AT 3007
                                                                                                                                                                                                                                                                                            3008 SSKATSSSSPRTATTLPVLTSTATKSTATSFTPIPSSTLGTTGTSONRPPHPMATMSTIH 3067
                                                                                                                                                                                                                                                                                                                                                                        3068 PSSTPETTHTSTVLTTKATTTRATSSMSTP---SSTPGTTWILTELTTAATTTAALPH-- 3122
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------GTPSSTPGTTWILTEPSTTATVTVPTGSTATASS------ 3156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3157 TRATAGTLKVLTSTATTPTVISSRATPSSSPGTA----TALP--ALRSTATTPTATSVT- 3209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3210 -----AIPSSSLGTAWTRLSQTTTP-----TATMSTATPSSTP----ETVHTSTV 3250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----TTRGSTVTPSSIPGTTHTATVLTTTTTVATGSMATPSSSTQTSGTPPSLTT 3362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3410 ---GTTHTPPVPNTTATTHGRSLPPSSP-----HTVP----TAWTSATSGILGTTH---- 3453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               480 YAELNPSIIVRAGSFDPSSQPNDRHVA-----FSTDGGKNWFQGSEPGGVT 525
                                                                                                                                                                          102 SSDQGATWQI------TPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                  197 KSSSSLGQASKTIFV-----GVADPNNPVFWSRDGGATW----QAVPGAPTGFIPHKG 245
                                                                                                                                                                                                                                                                                                                                                                                                                246 VFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTI 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             306 DRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       366 GVQPNPPVPSPKLG--WMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPM 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   424 VKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPS----TIFTSPVFTTGTSVD 479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   585 LSNGTFYRSTDGGVTFQPVAAGLPSSCAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSS 644
                                                                                                                                                                                                                                                      SGKGLWRST------DSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFD 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         526 TGGTVAASADGSRFVWAPG-DPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYA 584
                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metăzoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Submaxillary gland;
MEDLINE=91236743; PubMed=2033060;
BCkhard A.E., Timpte C.S., Abernethy J.L., Zhao Y., Hill R.L.;
"Porcine submaxillary mucin contains a cystine-rich, arboxyl-terminal domain in addition to a highly repetitive,
                                                                                                ; Score 154.5; DB 1; Length 5703;
; Pred. No. 1.2;
58; Mismatches 281; Indels 173;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---irepsr-----crshrpaarrcropsrpalss 3481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  645 WSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRS 686
A -> S (IN REF. 3).
DP -> RS (IN REF. 2).
F -> L (IN REF. 2).
A -> P (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-1989 (Rel. 12, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Apomucin (Mucin core protein) (Fragment)
                                                                                                3.8%;
                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
  601
629
633
676
                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sus scrofa (Pig).
                                                                                              Query Match
Best Local Simil
Matches 130; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APMU PIG
P12021;
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                        147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APMU_PIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FFFF
                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               . q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTM: EXTENSIVELY O-LINKED GLYCOSYLATED ON SER AND THR RESIDUES OF THE REPEAT UNITS. HIGHEST GLYCOSYLATION APPEARS TO OCCUR ON SER RESIDUES WHICH HAVE GLY AT POSITIONS AT +2 OR -2 FROM THE GLYCOSYLATION SITE OR, WHERE GLY IS THE PENULTIMATE RESIDUE. THE PERESENCE OF PROLINE (USUALLY AT POSITION +3 OR -3) APPEARS TO ALSO ENHANCE GLYCOSYLATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Gerken T.A., Owens C.L., Pasumarthy M.;

"Determination of the site-specific O-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:galnac transferase peptide binding site.";
J. Biol. Chem. 272:9709-9719(1997).
-!- FUNCTION: APOMUCIN IS PART OF MUCIN, THE MAJOR GLYCOPROTEIN STYNTHESIZED AND SECRETED BY MUCOUS CELLS OF THE SUBMAXILLARY GLAND. ITS HIGHLY VISCORS AQUEOUS SOLUTIONS SERVE TO LUBRICATE THE ORAL CAVITY AND TO PROTECT IT FROM THE EXTERNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY: Contains 1 VWFC domain.
SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBUNIT: INTERMOLECULAR DISULPIDE BONDS COULD HELP MAINTAIN A MULTIMERIC MUCIN STRUCTURE.
SUBCELLULAR LOCATION: Secreted.
TISSUE SPECIFICITY: SUBMATILIARY MUCOSAE.
DOMAIN: CONTAINS TANDEMLY REPEATED, IDENTICAL SEQUENCES OF 81
                                                                                                                                                                                                                                                                                                                    "Structural properties of porcine submaxillary gland apomucin.";
J. Biol. Chem. 262:11339-11344(1987).
                                                                           TISSUE=Submaxillary gland;
MEDLINE=88087170; PubMed=2826455;
MEDLINE=88087170; PubMed=2826455;
Timpte C.S., Eckhardt A.E., Abernethy J.L., Hill R.L.;
"Porcine submaxillary gland apomucin contains tandemly repeated, identical sequences of 81 residues.";
j. Biol. Chem. 263:1081-1088(1988).
                                                                                                                                                                                                                                    TISSUE-Submaxillary gland;
MEDLINE-87280220; PubMed-3611111;
Eckhardt A.E., Timpte C.S., Abernethy J.L., Toumadje A.,
Johnson W.C. Ofr., Hill R.L.;
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYDRATE-LINKAGE SITES, AND SEQUENCE OF 45-125.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                    Biol. Chem. 266:9678-9686(1991)
                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Submaxillary gland;
MEDLINE-97248516; PubMed-9092502;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M21174; AAA30990.1; -.
InterPro; IPR006208; Cys_Knot.
InterPro; IPR006207; Cys_Knot_C.
InterPro; IPR001007; VWF_C.
Pfam; PP00007; Cys_Knot; I.
SWART; SM00214; VWC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M61883; AAA30998.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS01185; CTCK 1; 1. PROSITE; PS01225; CTCK 2; 1. PROSITE; PS01208; VWFC 1; 1. PROSITE; PS50184; VWFC 2; 1.
                                                           SEQUENCE OF 1-503 FROM N.A
glycosylated domain.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Glycoprotein; Repeat.
NON TER 1
DOMAIN <1 36
                                                                                                                                                                                                                      SEQUENCE OF 45-80.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENVIRONMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESIDUES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
```

= -	449 TSIGRSGTTHTDLPGGTTIVLPGFSHSSQSSKPGSSVTTPGSPESGSETGTS	478 VDYAELNPSIIVRAGSFDPSSQPNDRHV	501 GEFSTIVISGSSHTEATTFIGGSGSPGTGSRPGTTGELSGTTIASGNATTEATTSTETRI	506 AFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPG	561 GPQT-GAQTTVPGSQVSGSETGTSEAVSNPAIASGSSSTGTTSGASDSQVTGSRTG	545 DPGQPVVYAVGFG-NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF	616	604 AAGLPSSGAVGVMFHAVPGKEGDLMLAASSGLYHSTNGGSSWSAITGVSSAVN VGFGKS :		065 - AFCSSIFANFVOSITGGVIGANTGATTGATTGATTANLLINDLANGGONG CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CO	708 QAITGDHANLRRYXIGT 724	: : 	PECIT.T 7		P21543; 01-MAZ1991 (Rel. 18, Created) 01-MAZ-1991 (Rel. 18, Last secuence update)		Alpha-amylase (EC 3.2.1.1)]. Paenibacilus polymyxa (Bacillus polymyxa).		SEQUENCE OF 1-936 STRAIN=72;		0 .			Isukagoshi N., Udaka S.; "A single gene directs synthesis of a precursor protein with beta- and alpha-amylase activities in Bacillus polymyxa."; J. Bacteriol. 171:375-382(1989).		<pre>MEDLINE=87231094; PubMed=2438660; Rhodes C., Strasser J., Friedberg F.;</pre>				polymyxa beta-amylase."; Biochemistry 30:4594-4599(1991).
	q _Q	ò	ব্য	ά	qu	à i	q	À da		ਨੇ _ਜ	ζ	qu	С	AMY	AC DT	DE TO	OOO	20 0 X	RP	RA RA	RT	RN RN	RA RA	RA RT RT RT RT RT RT RT RT RT RT RT RT RT	RN RP	RX	RT RI RI		R R R	Z Z
) (MUCIN TYPE).) (MUCIN TYPE).) (MUCIN TYPE).		(MUCIN TYPE). (MUCIN TYPE).			(MUCIN TYPE). (MUCIN TYPE). (MICIN TYPE).		(MUCIN TYPE). (MUCIN TYPE). (MUCIN TYPE).		MUCIN TYPE). MUCIN TYPE).		POTENTIAL).	(POTENTIAL). (POTENTIAL). (POTENTIAL)		1150; 3 194; Gaps 31;			ASIGQPETSRISVAGS- 293 GLTIDROHPNTIMVATOIS 320	SGASGST	LIFGVOPNPPVPSPKL 378	: SSGA8QAA 388		SKKTLEPGSYNTEAT 448
									: -: -	:::	???	; ; ;		:::::	777	:23: :::::	: :::	E E E			: 8	Indels	KS	I PHKGV	ASI SGLTID	SGTT-	AEPWI	GAPAV		7 1011
.0.	6 (INCOMPLETE). VWFC.	CTCK. BY SIMILARITY.	BY SIMILARITY. BY SIMILARITY.	222				O-LINKED (GALNAC O-LINKED (GALNAC O-LINKED (GALNAC					O-LINKED (GALNAC) O-LINKED (GALNAC) O-LINKED (GALNAC)		O-LINKED (GALNAC) () O-LINKED (GALNAC) () O-LINKED (GALNAC) ()	(GALNAC)	777	(GALNAC) ((GALNAC) ((GLCNAC) ((GLCNAC) (N-LINKED (GLCNAC) (PC N-LINKED (GLCNAC) (PC N-LINKED (GLCNAC) (PC	MW; 3CB68B5D29DD7F5A CR	Pred. No. 0.21; 75; Mismatches 261;	TYIANPTDTTGYQSDIQGVVWVAFDKS	TO VAGOOGAR	FTTGTVSGASGSTGSSSGSPGATGASI SVTSGTWTRISPVPSTDTANDYFGYSGLTID	: :	SATWTRIWDWTSYPNRSLRYVLDISAEPWI		WMLYGTGATLY-ATNDLTK	averna en ar reenea arveanaigilea Va
368	391 6 995 VW	062 1146 CT	1123 BY	089 1141 BY SIMILAI ? 1145 BY SIMILAI	46 O-LINKED 50 O-LINKED	51 O-LINKED 57 O-LINKED 69	61 O-LINKED (66 O-LINKED (67 O-LINKED (73 O-LINKED (74 O-LINKED (O-LINKED O-LINKED O-LINKED	83 O-LINKED (93 O-LINKED	94 O-LINKED 096 O-LINKED 098 O-LINKED 0	101 O-LINKED (104 O-LINKED (GALNAC) O 106 O-LINKED (GALNAC) O 107 O-LINKED (GALNAC)	108 O-LINKED (GALNAC) 10 O-LINKED (GALNAC)	114 O-LINKED (GALNAC) (117 O-LINKED (GALNAC) (123 O-LINKED (CALNAC)	123 O-LINKED (GALNAC) (124 O-LINKED (GALNAC) (418 N-LINKED (GLCNAC) (547 N-LINKED (GLCNAC) (917 N-LINKED (GLCNAC) ((GLCNAC) (GLCNAC)	AA; 109615 MW; 3CB68B5D29DD7F5A CR	Pred. No. 0.21; 75; Mismatches 261;	SGATWSQMTNPPDVGTYIANPTDTTGYQSDIQGVVWVAFDKS	GOFGALGAS LGQFELON LOVAGOS GAPENS CAPENS TETARPSVAGSGTTGTVSGASGSTGSSSGSPGATGASIGQPETSR TGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDXFGYSGLTIDROHPNTI		WWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKL	SGSPGATGASI	GWMDEAMAIDPFNSDRMLYGTGATLY-ATNDLTKWDSGGQIHIAP 	. con a construction of the part of the part of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second	

3 8 3 8 3 8 3 8 3 8 3

```
1196 AA; 130893 MW; A41EA6B70F257064 CRC64;
                                                                                  Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :||: | | ||
723 SGAFTSMDLGT 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         681 TGAYRSDDCGT 691
                                                               Similarity
                                                                                                                                                                                                                                                                                                                                  144
  SEQUENCE
                                          Query Match
                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PGCA_MOUSE
                                                               Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEEPPRE
                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                        CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-glucosidic linkages in
                                                           polysaccharides so as to remove successive maltose units from the non-reducing ends of the chains.

CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides and polysaccharides.

SUBCELLULAR LOCATION: Secreted.

SIMILARITY: In the N-terminal section; belongs to family 14 of glycosyl hydrolases.

SIMILARITY: In the C-terminal section; belongs to family 13 of
  PRODUCE MULTIFORM BETA-AMYLASES AND A 48 KDa ALPHA-AMYLASE AFTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C-SS: SEPLED DECREASE IN ACTIVITY.
C-SS: 60-POLD DECREASE IN ACTIVITY.
C-SS: 60-POLD DECREASE IN ACTIVITY.
N -> S: 60-POLD DECREASE IN ACTIVITY.
N -> S: IN REF: 3).
N -> D: IN REF: 3).
N -> D: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> T: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
N -> C: IN REF: 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00642; Aamy; 1.
SMART; SM00632; Aamy C; 1.
PROSITE; PS00506; BETA AMYLASE 1; 1.
PROSITE; PS00679; BETA AMYLASE 2; 1.
Multifunctional enzyme; Hydrolase; Glycosidase; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BETA/ALPHA-AMYLASE.
BETA-AMYLASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALPHA-AMYLASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR006589; Alp amyl cat sub. InterPro; IPR006048; Alpha_amyl_C. InterPro; IPR006048; Alpha_amyl_cat. InterPro; IPR006046; Alpha_amyl_cat. InterPro; IPR005085; CBM 25. InterPro; IPR005085; Glyco_hydro_l3. InterPro; IPR005086; Glyco_hydro_l4. Pfam; PF00128; alpha-amylase.; I. Pfam; PF01373; CBM 25; 2. Pfam; PF01373; CBM 25; 2. Pfam; PF01373; Glyco_hydro_l4; 1. PRINTS; PR00110; ALPHAAMYLASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polysaccharide degradation; Repeat.
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M15817; AAA85446.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, Y00150, CAA68344.1; -. PIR, A29130, A29130.
HSSP; P36924; 1B9Z.
                                                                                                                                                                                                                            glycosyl hydrolases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                35
1196
454
454
558
668
1196
126
198
394
318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
DISULFID
ACT SITE
ACT SITE
MUTAGEN
MUTAGEN
MUTAGEN
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
```

```
349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | : | : | : | | : | | 3.40 LDXINA-AWGTKLTS--LSQINPPTDGDGFYTNGGY------NSAYGKDFLSWYQS 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                390 RLNGENALPTGGSGFQKIEEKITKFGYHGFTLLRINNLVN-NDGSP----TGELSGFKQY 444
                                                                                                                                                                    61 NNWG------YNGVVSIAAD-----PINTNKV-WAAVGMYTNSWDPNDGAILRS 102
                                                                                                                                                                                                                                            53 NDWGSFKKQLQTLKNNGVYAITTDVWWGYVESAGDNQFDWSYYKTYANA-----V 102
                                                                                                                                                                                                                                                                                                                                 103 SDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWS 162
                                                                                                                                                                                                                                                                                                                                                                               264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               265 YDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPD 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            343 LIQKFKDADLD--LTFTCLEMSDSGTAPNYSLPS-----TLVDTVSS--IANAKGV 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              190 RAGSFDPSSQPNDRHVAFSTDGGKNWFQGS-------EPGGVTTGGTVAASADG 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          505 DIGS-----ASQLEAAFN-DGNNNWDSNNTKNYSFSTGTSTYTPGNSGNAGTITSGAPA 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           558 GA---NPGDGGGTTNKVTVYÝKKGFNSPYIHYRPAGGSWTAAPGVKMQDAEISGYAKITV 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   615 DIGSASQLEAAFNDGNNNWDSNNTKNYLFSTGTSTYTP----GSNGAAGTIRTGAP--S 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       625 GDLWLAASSGLY----HSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGV 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    668 GSV-LSVVTSTYATDLNEVTGPIQTEKLSGVS--LNVSTSTYAPNSN--GVEVTAQTEAP 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            350 SLRYVLDISAEPWLTF-----GVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGA 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         445 IISKAKPDNNGGTGNKVTIYYKKGFNSPYIHYRPAGGSWTAAPGVKMQDAEISGYAKITV 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           575 DRVNPKTFYAL-----SNGT-PYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKE 624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PGCA_MOUSE STANDARD; PRT; 2132 AA.

Q61282; Q64021;
Q1-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
15-MAR-2004 (Rel. 43, Last annotation update)
protein) (CSPCP).
AGCI OR AGC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          537 SRFVWAPGDPG-----QPVVYAVGF-----GNSWAASQGVPAN----AQIRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         163 OMTNFPDVGTYIANPTDTTGY-OSDIOGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OF------KDESGYANSEALSPLWSGTGKQYDELYASFAENFAGYKSIIPKIY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              222 WSRDGGATWQAVPGAPTGFI------PHKGVF----DPVNHVLYIATSNTGGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          191 LS------GGPSGELRYPSYYPAAGWSYPGRGKFQAYTETAKNAFRTAMNDKYGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               325 TI-----IMDWTSYPNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           287 VLEKHLGVIGAAAHKNFDSVFGVRIGAKISGLHWQMNNPAMPHGTEQAGGYYDY----NR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            102 TLYATNDLTKWDSGGQIHIAPMVK----GLEETAVNDLISPPSGAPLISALGDLGGF---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---THADVTAVPSTIFTSPVF-------PSIIV
3.8%; Score 152; DB 1; Length 1196;
21.1%; Pred. No. 0.25;
Live 72; Mismatches 282; Indels 270; Gaps
```

```
ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as Iong as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         anu us.

-I-PTM: Contains mostly chondroitin sulfate, but also keratan sulfate chains, N-linked and O-linked oligosaccharides.
-I-DISEASE: Defects in AGC1 are the cause of cartilage matrix deficiency (CMD). CMD is an autosomal recessive syndrome characterized by cleft palate, short limbs, tail and snout.

Mutation in strain CMD causes absence of aggrecan by truncation of the protein (mutation in mimunoglobulin-like V-type domain.
-I-SIMILARITY: Contains 1 limmunoglobulin-like V-type domain.
-I-SIMILARITY: Contains 1 Loft-like domain.
-I-SIMILARITY: Contains 1 Sushi (SCR) domain.
-I-SIMILARITY: Contains 1 Sushi (SCR) domain.
-I-SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             similarity).

DOMAIN: Two globular domains, Gl and G2, comprise the amino terminus of the proteoglycan, while another globular region, G3, makes up the COOH terminus. Gl contains link domains and thus consists of three disulfide-bonded loop structures designated as the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS) and the chondroitin sulfate (CS) attachment domains lie between G2 and G3.
                                                                                                                                                                                                                                                                                                                                                                                                                    J. Biol. Chem. 274:2044-20449(1999).

-!- FUNCTION: This proteoglycan is a major component of extracellular matrix of cartilagenous tisues. A major function of this protein is to resist compression in cartilage. It binds avidly to hyaluronic acid via an amino-terminal globular region. May play a regulatory role in the matrix assembly of the cartilage.
-!- SUBONIT: Interacts with FBLN1.
-!- SUBORIL!! ACCATION: Secreted; extracellular matrix (By
                                                                                                                                                                                                                                                                      Yamada Y.; "Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
                                                                                                                                                                                                                                                                                                                                             INTERACTION WITH FBLM1.
MEDLINE=99329059; PubMed=10400671;
Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
"Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                          Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
                                                                                                                                              "Complete coding sequence, deduced primary structure, chromosomal localization, and structural analysis of murine aggrecan."; Genomics 22:364-371(1994).
                                                                                         STRAIN=BALB/c; TISSUE=Cartilage;
BEDINE=25104847; PubMed=7866222;
WALCE E., Deak F., Erhardt P., Coulter S.N., Fueloep C., Horvath
Doege K.J., Glant T.T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; AS5182; AS5182.
HSSP; P98066; 1TSG.
MGD; MGI:99602; Agc1.
InterPro; IPR002353; Antifreezell.
                                                                                                                                                                                                                                         MEDLINE=95004579; PubMed=7920633;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S73721; AAB32160.1; JOINED
                                                                                                                                                                                             [2]
SEQUENCE OF 211-326 FROM N.A.
STRAIN=129/Sv;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; L07049; AAC37670.1; -. EMBL; S73722; AAB32160.1; -.
                                                                                                                                                                                                                                                                                                   the aggrecan gene.";
Nat. Genet. 7:154-157(1994).
                                           NCBI_TaxID=10090;
                                                                           SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                            versican."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZMBL;
```

```
64 GYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGA-----TWQITPL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 0.74;
96; Mismatches 286; Indels 292; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           n 3.7%; Score 149; DB 1; Length 2132; Similarity 19.5%; Pred. No. 0.74;
                                                                                                                                                                                                                                                                  PROSITE; PS00615; C TYPE LECTIN 1; 1.
PROSITE; PS00615; C TYPE LECTIN 2; 1.
PROSITE; PS002091; C TYPE LECTIN 2; 1.
PROSITE; PS002091; IG MFG; 1.
PROSITE; PS01241; LINK; 4.
Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OB2BCDFC6CBDA163 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                          AGGRECAN CORE PROTEIN IG-LIKE V-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELL ATTACHMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY LINKED (GLCUN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C-TYPE LECTIN.
SUSHI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
InterPro; IPR003129, 194 The InterPro; IPR001306; Ig VinterPro; IPR001304; Lectin_C.
InterPro; IPR001304; Lectin_C.
InterPro; IPR001304; Lectin_C.
InterPro; IPR0003324; SGXXSG.
InterPro; IPR000436; Sushi_SCR_CCP.
Pfan; PP00047; ig; 1.
Pfan; PP00193; SGXXSG; 60.
Pfan; PP00193; SGXXSG; 60.
Pfan; PP00193; Xiink; 4.
PRINTS; PR001365; ANTIFREEZEII.
PRINTS; PR001365; ANTIFREEZEII.
PROMO PP000918; Link; 4.
SWART; SW00031; CCP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                        LINK 2.
LINK 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G1-A.
G1-B.
G1-B'.
G2-B.
G2-B'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KS.
CS-1.
CS-2.
G3.
                                                                                                                                                                                                                                                                                                                                                                                                                         LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2132 AA; 222008 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 163; Conservative
                                                                                                                                                                                                                                          SMART; SM00406; IGV; 1.
SMART; SM00445; LINK; 4.
PROSITE; PS00615; C_TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1933
2042
2034
2034
2105
126
239
333
333
                                                                                                                                                                                                                                                                                                                                                                  Immunoglobulin domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1922
1950
2018
2049
2078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
```

39,

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       resolution.";
J. Mol. Biol. 281:885-897(1998).
-!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
linkages in glycogen, amylopectin and their beta-limits dextrins.
-!- COFACTOR: Binds I calcium ion per subunit.
-!- SUBUNIT: Monomer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A -> G (IN REF. 1).
FA -> C (IN REF. 1).
L -> V (IN REF. 1).
L -> V (IN REF. 1).
GAYT -> AVH (IN REF. 1).
SGLDIFARPWAGGGNSVOUGHFDS ->
TVWICLENLIGPSAATRISWVDSRRVVRVEWSVPRQ (IN
                                                                                                                                               Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H., "Nucleotide sequence and expression of the isoamylase gene from an isoamylase-hyperproducing mutant, Pseudomonas amyloderamosa JD210."; Biochim. Biophys. Acta 1087:309-315(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;
"Three-dimensional structure of Pseudomonas isoamylase at 2.2-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PDB; 1BF2; 12-AUG-98.
InterPro; IPR006047; Alpha_amyl_cat.
InterPro; IPR004193; G1yco_hydro_l3N.
InterPro; IPR001103; G1yco_hydro_l3N.
InterPro; IPR007110; Ig-like.
Pfam; PF02922; isoamylase; 1.
Hydrolase; G1ycosidase; N; 1.
SiGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REF. 1).
WP -> S (IN REF. 1).
AFRKAHPA -> RSARHIP (IN REF. 1)
"Cloning and nucleotide sequence of the isoamylase gene from Pseudomonas amyloderamosa SB-15."; J. Biol. Chem. 263:9271-9275(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -i- INDUCTION: By maltose.
                                                                                                                                                                                                                                                                                                                             Amemura A., Fujita M., Futai M.;
"Transcription of the isoamylase gene (iam) in Pseudomonas
amyloderamosa SB-15.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (VIA CARBONYL OXYGEN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISOAMYLASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CALCIUM.
CALCIUM.
CALCIUM.
                                                                                                                                                                                                                                                                                                                                                                                                                                     K-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CALCIUM
                                                                                                                                                                                                                                                                                                                                                                                         Bacteriol, 171:4320-4325(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=98387895; PubMed=9719642;
                                                                                                                                                                                                                                                                                                       MEDLINE=89327147; PubMed=2753857;
                                                                                                                                  MEDLINE=91064385; PubMed=2248978;
                                                                                                                                                                                                                                                           SEQUENCE OF 744-776 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; J03871; AAA25854.1; -. EMBL; X13378; CAA31754.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26
776
401
461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   536
154
255
255
256
258
258
422
422
422
422
422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       556
                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4461
1536
2555
2556
2556
2556
2556
7386
7386
4113
454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       555
650
                                                                                                                                                                                                                                                                                   STRAIN=SB-15;
                                                                                                              STRAIN=JD210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               METAL
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            METAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     METAL
    : | | | | | | | : 1 | : | | | : | : | | | : | | : | | | : | | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GYVSGIPSGGDGTET----SASGVEDVSGLPSGGEGLETSASGVEDLGPSTRDSLETSAS 1160
                                                                                                   .161 GVDVTGFPSGRGDPETSVSGVGDD---FSGLPSGKEGLETSASGA--EDLSGLPSGKEDL 1215
                                                                                                                                                                                  .591 EAGEGPSGILEFSGAHSGTPDISGELSGSLDLSTLQSGQMETSTETPSSPY--FSGDFSS 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1649 TTDVSGESIAATTGSGESSGLPEVTLNTSELVEGVTEPTVSQELGHGPSMTYISRLSEAS 1708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1709 GDASASGDLGGAVTNFPGSGVEASVPEASSDLSAYPEAGVGVSAAPEASSKLSEFPDLHG 1768
                                                                                                                                                                                                                                                                    1307 GSGEV-----SGL----PLGELDSSAD---ISGL-----PSGTELSGQASGSPDS-- 1344
                                                                                                                                                                                                                                                                                                                                                                                                                 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1345 SGETSG------BETSGI 1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                553 AVGFGNSW----AASQGVP-----ANAQIRSDRVNPKTFYALSNGTFYR 592
                                                                                                                                                        ------VGTYIANPT-DTTGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVA 214
                                                                                                                                                                                                                                                                                                                             268 SSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTII 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   388 DPFNSDRMLYGTGATLYATNDLTKWD--SGGQIHIAPMVKGL-------EET 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            517 QGSEPG------GVTTGGTVAASA-DGSRFVWAPGDP-------GQPVVY 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --STNGGSSWSAIT----GVSSAV-----GVSSAV-----GVGSSAVFVVGTIGG 679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1769 ITSAFHETDLEMTTPSTEVNSNPWTF--QEGTREGSAAPEVSGESSTTSDIDTGTSG 1823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          680 VTGAYRSDDCGTT------WVLINDDQHQYGNWGQAITGDHANLRRVYIGTNG 726
                                                                                                                                                                                                                                                                                                                                                                                                                 328 FRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAI
                                                                  115 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPD----
                                                                                                                                                                                                                                       DPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       431 AVNDLISPPSGAPL-----ISALGDLGGFTHADVTAVPSTIFT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----SPVFTTGTSVDYAELNPSIIVRAGSFDPSSOPNDRHVAFSTDGGKNWF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  593 STD-GCVTFQPVAAGLPSSGAVGVMPHAVPGKEGDLWLAASSGLYH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=SB-15;
MEDLINE=88243808; PubMed=3379068;
Amemura A., Chakraborty R., Fujita M., Noumi T., Futai M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-WAR-1989 (Rel. 10, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            776 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (soamylase precursor (EC 3.2.1.68)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pseudomonas amyloderamosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria, Proteobacteria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MCBI_TaxID=32043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ISOA PSEAY
P10342;
                                                                                                                                                                                                                                          215
                                                                                                                                                                                                                                                                                   1267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
ISOA PSEAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                       셤
                                                                                                                                                 à
                                                                                                                                                                                           셤
                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

FT HELIX FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TU	Query Match 3.7%; Score 148.5; DB 1; Length 776; Best Local Similarity 20.7%; Pred. No. 0.24; Matches 171; Conservative 88; Mismatches 275; Indels 291; Gaps 47;	QY 27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86	QY 87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDFNNDNILYFGAP 146	QY 147 SGKGLMRSTDSGATWSQMTNFPDVGTYIANPTD-TTGYQSDIQGVVWVAFDKSSSS 201	Qy 202 LGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTG-F 240	Qy 241 IPHKGVFDPVNHVLXIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST 291	QY 292DTANDYF-GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIW 341	QY 342 DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAWAIDPFNSDRML 396	397 YGTGATLYATNDLTKWDSGGOIH-IAPMVKGLEETAVNDLISPPSGAPLI
•									
								·	
									•
31 440 440 440 440 440 440 440 44	2 2 2 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	270 278 284 288	3310 312 312 313	328 339 351	354 357 361	371 374 392 392	4 4 4 4 002 708 80	421 424 29	432 435
334 334 336 337 338 338 338 338 338 338 338 338 338	255 255 259 265 265	269 278 284 286	254 311 314	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	352 354 361	371 373 392	397 401 403	4 4 4 4 4 4 4 4 4 4 2 3 3 3 3 3 3 3 3 3	431 434
FT TURN FT STRAND FT STRAND FT STRAND FT STRAND FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TURN FT TUR		FT TORN FT STRAND FT STRAND FT HELIX						•	

47;

291;

146

201

147

98

```
148 AQEVSQDP----LNPSNQ-----NGNVFASGASYR-----TTDSGIY--AP 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DWTS-YPNRSLRYVLDISAEPWLTFGVOPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR-------AGS 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              349 TTYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR----F
                                                                                                                                                                                                                                                                                                 GAPGILYVRIDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPININKVWAAVG
                                                                                                                                                                                                                                                                                                                              GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY
                                                                                                                                                                                                                                                                                                                                                                                                                                           243 IG-VTAVEFLPVOETONDANDVVPNSDANONYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QAMVQAFHNAGIKVYMDV-------VYNHTAEGGTWTSSDPTTATIYSWRGLDN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DLASVLGNSCLNGAYTASA------PNCPNGGYNFDAADSNVAINRILREFTVRPAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503 YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                494 FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRF-VWAPGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---SWAASQGVPANAQIRSDR-----V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       602 PLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQ---SNFYTFAQRLIAFRKAHPAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     578 NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVG------VMFHAVPGKEGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LWLAASSGLYHSTNG-----GSSWSAIT-----GVSSAVNVGFGKSAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --LGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDTCDWNDGASTFV----APG
                                                                                                                                                                                                                                                                                                                                                             87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                                                                         SGKGLWRSTDSGATWSQMTNPPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LGQASKTIFVGVAD-----PN---PN---NPVFW-----SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     397 YGTGATLYATNDLTKWDSGGQIH-IAPWVKGLEE------TAVNDLISPPSGAPLI
                                                                                            CALCIUM (VIA CARBONYL OXYGEN) (BY
                                                                                                                                                                                                                                    Length 776;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               666 SSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710
ISOAMYLASE.
BY SIMILARITY.
BY SIMILARITY.
CALCIUM (BY SIMILARITY).
CALCIUM (BY SIMILARITY).
                                                                                                                                                                                                      A5E4C02EF026A3A4 CRC64;
                                                                                                         SIMILARITY).
CALCIUM (BY SIMILARITY)
CALCIUM (BY SIMILARITY)
                                                                                                                                                                                                                                   3.7%; Score 148.5; DB 1;
20.7%; Pred. No. 0.24;
iive 88; Mismatches 275;
                                                                                                                                                      BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           546 P----GQPVVYAVGFGN------
                                                                                                                                                                                                        83656 MW;
                                                                                                                                                                                                                                                                 171; Conservative
                                                                                                                         258
285
422
616
766
   776
401
461
536
154
255
                                                                                                                        258
285
410
546
738
776 AA;
                                                                                                                                                                                                                                                     Similarity
 27
4401
461
536
154
255
CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  701
                                                                                                                                                         DISULFID
                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                      Local
                                                                            METAL
METAL
                                                                                                                           METAL
                                                                                                                                                                                                                                                      Best Loca
Matches
                                                              METAL
                                                                                                                                            METAL
   STATESTA
                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the Buropean Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce).
                                                                                                                                                                         577
                                                                                                                                                                                                                                                                   700
                                                                                                                                                                                                                                                                                                 665
 493
                                                                                                                                           601
                                                                                                                                                                                                                                                                                                                                751
                                                                   P----GOPVVYAVGFGN------SWAASOGVPANAQIRSDR-----V
                                                                                                                                                                                                                                                        --LGDSNSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTDTCDWNDGASTFV-----APG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of Pecudomonas sp.",
J. Gen. Microbiol. 135:37-45(1989).
-!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1,6)-D-glucosidic branch
linkages in glycogen, amylopectin and their beta-limits dextrins.
-!- COFACTOR: Binds I calcium ion per subunit (By similarity).
-!- SUBUNIT: Monomer (By similarity).
-!- INDUCTION: By maltose.
                                                SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR-------AGS
                                                                                                            494 FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRF-VWAPGD
                                                                                                                                           -----DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEMLSAGT
                                                                                                                                                                                                      PLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQ---SNFYTFAQRLIAFRKAHPAL
                                                                                                                                                                                                                                    578 NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVG------VMFHAVPGKEGD
                                                                                                                                                                                                                                                                                                 LWLAASSGLYHSTNG------GSSWSAIT-----GVSSAVNVGFGKSAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Cloning and nucleotide sequence of the isoamylase gene from a strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
                                                                                                                                                                                                                                                                                                                                                              SSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710
                                                                                                                                                                                                                                                                                                                                                                                             -----YGOCGOSL 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             rognoni A., Carrera P., Galli G., Lucchese G., Camerini
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase, Glycosidase, Calcium-binding, Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P26501,
01-AUG-1992 (Rel. 23, Created)
101-AUG-1992 (Rel. 23, Last sequence update)
101-AUR-2004 (Rel. 43, Last annotation update)
Isoamylase precursor (EC 3.2.1.68).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        776 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.
MEDLINE=89381677; Pubmed=2778432;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                             SE-----ALIGG-----AGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hydro_13N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPR006047; Alpha amyl cat.
Interpro; IPR004193; Glyco hydro_13N
Interpro; IPR007110; Ig-11ke.
Pfan; PF00128; alpha-amylase_1.
Pfan; PF02922; isoamylase_N; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pseudomonas sp. (strain SMP1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M25247; AAA25855.1; -.
EMBL; A10909; CAA00929.1; -.
HSSP; P10342; 1BF2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacteria, Proteobacteria.
NCBI_TaxID=306;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                            YGPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISOA PSESP
                                                                                                                                           557
                                                                                                                                                                           546
                                                                                                                                                                                                                                                                                                                                                              999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Grandi
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ISOA PSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                         셤
                                                                                                        8
                                                                                                                                    g
                                                                                                                                                                        ò
                                                                                                                                                                                                g
                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                      음
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                      셤
```

400

445 502

451

577 658 626 751

601

348

301 291

```
CONFLICT
CONFLICT
SEQUENCE
 REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
CONFLICT
                                                                                                                                                                                                                              44
                                                                                                                                                                                                                                                                          95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           535
                                                                                                                      CONFLICT
                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
  g
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Busse and institute of Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAUTION: Ref.3 sequence differs from that shown due to frameshifts in positions 414 and 732.
                                                                                                                                                                                                                                                                                            MEDLINE=21534948; PubMed=11677609; McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Coutrney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D. Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
                                                                                                                                                                                                                                                                                                                                                                    "Complete genome sequence of Salmonella enterica serovar Typhimurium
                                                                                                                                                                                                                                                                                                                                                                                                                                                               "High-level expression of Escherichia coli NADPH-sulfite reductase: requirement for a cloned cysG plasmid to overcome limiting siroheme
                                                                                                                                                     Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIĀL.
PUTATIVE SURFACE-EXPOSED VIRULENCE
PROTEIN BIGA.
 -----TLIGG-----AGTT-----YGQCGQSL 770
                                                                                                                    Putative surface-exposed virulence protein bigA precursor
                                                                                                                                                                                                                        Stojiljkovic I., Valentine P., Heffron F.;
"Salmonella typhimurium rhs homolog.";
Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15 X 11 AA TANDEM REPEATS.
1 (INCOMPLETE).
2 (INCOMPLETE).
3 (INCOMPLETE).
                                                        BIGA_SALTY STANDARD; PRT; 1953 AA. P25927; P25928; Q9XCQ3; C1-MAY-1992 (Rel. 22. Ceated) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Complete proteome,
                                                                                                                                                                                                                                                                                     STRAIN=LT2 / SGSC1412 / ATCC 700720;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M64606; AAA27042.1; ALT FRAME.
EMBL; M64606; AAA27043.1; ALT FRAME.
                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=91100301; PubMed=1987123; Wu J.Y., Siegel L.M., Kredich N.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4264596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J. Bacteriol. 173:325-333(1991)
-!- CAUTION: Ref.3 Rominand Aif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF133696; AAD39458.1; -.
EMBL; AE008859; AAL22340.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-765 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Virulence, Repeat, Signal, SIGNAL 1 27 CHAIN 28 1953
                                                                                                                                                                                                                                                                                                                                                                                           Nature 413:852-856(2001).
                                                                                                                                          Salmonella typhimurium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  StyGene; SG10437; bigA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C39200
                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                               STRAIN=ATCC 14028;
                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A
                                                                                                                                                                             NCBI_TaxID=602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           101
101
101
101
104
1134
1145
1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C39200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D39200
                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=LT2;
                                                                                                                                BIGA OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RPEAT
                                              SALTY
셤
```

```
1398 NSGIINVGTEQGK-------EDGTNGTGLIGIKGNGNATTINNTADGVINVYADDSYA 1448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1532 TLKANNLVIGDNVKVDTGFTSGTADTTVVVDNAFTGSNIQGADNIT-----STS 1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1344 WNTATGVINGINPDAPLINLGRGYNFGNQGTINVQGD----NAV--AISGGTSSYVINLV 1397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1787 IADSDMRQQYLEFRSEGAKTFTWMGDALKVTPYAGVKFRHTMEDGYKERSAGDFNLSMNS 1846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1847 GNETAVDSIVGLKLDYAGKDGWSATATLEGGPNLSYSKSQRTASLQGAAGOSF-GVDDGQ 1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                211 VGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPYDGSSG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NDGAILRSSDQGATWQITPLPFKLGGNMPG----RGMGERLAVDPNNDNILYFGAPSGKG 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         151 LWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIF 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        271 DVWKFSV-----TSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQ 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 319 ISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKL 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       475 GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA 534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DGSRFVWAPGDPGQPVVYAVGFG--NSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYR 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                634 ------GLYHSTNGGSSWSAITGVSSAVNVGFGKS------APGSSYPAVFVVG 675
                                                                                                                                                                                                                                                                                                                                                                                                                                         WDAANG-----RWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWD---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --IHIAPMVK-GLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GGVTFQPVAAGLPSSGAVGVMF-HAVPG--KE---GDLWLAASS
                                                                                                                                                                                                                                                                                                                                                                                             257; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1906 KGGGVNGLATIGVKYSSND----TALHLDAYQWKEDGISDKGFMLNVKKTF 1952
                                                                                                                                                                                                                                                                                                                                                   Length 1953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         676 TIGGVTG-----AYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVY
                                                                                                                                                                                      (IN REF. 1)
                                                                                                                                                                                                                                                                                                                                                                                               Indels
8.

9.

11.

12.

13.

14.

15 (INCOMPLETE).

D -> DRGDDDVTPPDD (IN REF. A -> R (IN REF. 3).

D -> N (IN REF. 1).

QYLE -> ITLQ (IN REF. 1).

SA -> T (IN REF. 1).
                                                                                                                                                                                                                                                                                                    AA; 200150 MW; 611B3F1C954D91AE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        379 GWMDEAMAIDPFNSDRMLY----GTGATLYATNDLTKWDSGQ-
                                                                                                                                                                                                                                                                                                                                                                                           Matches 158; Conservative 83; Mismatches 274;
                                                                                                                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                                                                                                                                                                                              3.7%; Score 148.5; Di
20.5%; Pred. No. 0.71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1717 NLTLEY------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1506 TPPADPNAPQOLSN------
    177
188
199
210
221
232
243
252
252
252
207
252
207
214
1698
11798
                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                           514
1698
1795
1836
                                                                                                                                                                                                                                                                                                    1953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         593 STD----
```

```
MEDLINE-86259736, PubMed=3460082;

MEDLINE-86259736, PubMed=3460082;

Sai S., Tankar T., Kobher R.A., Tanzer M.L.;

"Cloning and sequence analysis of a partial CDNA for chicken cartilage proteoglycan core protein.";

Proc. Natl. Acad. Sci. U.S.A. 83:5081-5085(1986).

-!- FINCTION: This proteoglycan is a major component of extracellular matrix of cartilagenous tissues. A major function of this protein is to resist compression in cartilage. It binds avially to hyaluronic acid via an amino-terminal globular region. May play a regulatory role in the matrix assembly of the cartilage.
-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
PGCA_CHICK STANDARD; PRT; 2109 AA.
P07858; Q90810; Q90820; Q90991; Q91047;
01-AUG-1988 (Rel. 08, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Aggredan core protein precursor (Cartilage-specific proteoglycan core
                                                                                                                                                         Gallus gallus (Chicken).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Archosauria, Aves, Neognathae, Galliformes, Phasianidae, Phasianinae,
                                                                                                                                                                                                                                                                                                    MEDLINE=94043149; PubMed=8226878;
Li H., Schwartz N.B., Vertel B.M.;
"CDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia.";
J. Biol. Chem. 268:23504-23511(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE-Embryo;
MEDLINE=90307744; PubMed=1694853;
Krueger R.C. Jr., Fields T.A., Mensch J.R. Jr., Schwartz N.B.;
"Chick cartilage chondroitin sulfate proteoglycan core protein. II.
Nucleotide sequence of CDNA clone and localization of the S103L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1894-2109 FROM N.A.
SPEDILINES-89008500, PubMed=3170613;
Tanaka T., Har-El R., Tanzer M.L.;
"Partial structure of the gene for chicken cartilage proteoglycan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chandrasekaran L., Tanzer M.L.;
"Molecular cloning of chicken aggrecan. Structural analyses.";
Biochem. J. 288:903-910(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN=White leghorn; TISSUE=Chondrocytes;
MEDLINE=95128519; PubMed=7827752;
Primorac D., Stove M.L., Clark, Rowe D.W.;
"Molecular basis of nanomelia, a heritable chondrodystrophy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1-1855 AND 1893-2109 FROM N.A.
                                                                                                                                                                                                                                                                                           STRAIN=White leghorn; TISSUE=Embryo;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     epitope.";
J. Biol. Chem. 265:12088-12097(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               core protein.";
J. Biol. Chem. 263:15831-15835(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Cartilage;
MEDLINE=93111968; PubMed=1339285;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=94107258; PubMed=8280087;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chandrasekaran L., Tanzer M.L.;
Biochem. J. 296:885-887 (1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1042-1559 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1492-1610 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matrix Biol. 14:297-305(1994).
                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                   NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SRRATUM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 chicken
```

similarity).

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                     -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate chains, N-linked and O-linked oligosaccharides.
-!- DISEASE: DEFECTS IN AGCI ARE THE CAUSE OF NANOMELIA, A LETHAL CONNECTIVE IN AGCI ARE THE CAUSE OF NANOMELIA, A LETHAL CONNECTIVE TISSUE DISORDER APPECTING CARTILAGE DEVELOPMENT INTECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT LIMBS.

(CHONDRODYSTROPHY) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS. AGGRECAN IS TRUNCATED AT ITS C-TERMINAL IN THE CS-2 BINDING DOMAIN AND IS NOT ANYMORE SECRETED FROM THE CHONDROCYTES.
-!- SIMILARITY: Contains 1 immunoglobulin-like v-type domain.
-!- SIMILARITY: Contains 1 immunoglobulin-like domain.
-!- SIMILARITY: Contains 1 C-type lectin family domain.
-!- SIMILARITY: Contains 1 C-type lectin family domain.
                                                                               Isoid=P07898-2; Sequence=VSP 003073; DOMAIN: Two globular domains, Gl and G2, comprise the amino terminus of the proteoglycan, while another globular region, G3, makes up the COOH terminus. Gl contains link domains and thus consists of three disulfide-bonded loop structures designated as the A, B, B' motife. G2 is similar to G1. The keratan sulfate (KS) and the chondroitin sulfate (CS) attachment domains lie between G2
                  Event=Alternative splicing; Named isoforms=2;
                                                  IsoId=P07898-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SWART; SM00032; CCP; 1.
SWART; SM00034; CLECT; 1.
SWART; SM001034; EGF_CA; 1.
SWART; SM00409; IG; 1.
SWART; SM00445; LINK; 4.
FROSITE; PS00010; ASX HYDROXYL; 1.
PROSITE; PS000115; C_TYPE_LECTIN_1; 1.
PROSITE; PS00012; EGF_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M88101; -; NOT ANNOTATED_CDS.
EMBL; S74657; AAC60751.1; -.
EMBL; S74656; AAC60751.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002353; AntifreezeII
InterPro; IPR000152; Asx_hydroxyl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00008; EGF; 1.
Pfam; PF00004; ig; 1.
Pfam; PF00147; ig; 1.
Pfam; PF02139; SGXXSG; 56.
Pfam; PF00193; Xiink; 4.
PRINTS; PR001265; ANTIFREZEII.
PRINTS; PR01265; ININMODULE.
ProDom; PD000918; Link; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; S74657; AAC60751.1; -...
EMBL; S74656; AAC60751.1; JOINI
EMBL; J04028, AAA48719.1; -..
EMBL; M13933; AAA48720.1; -..
PIR; I50421; I50421.
HSSP; P08709; IBF9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; L21913; AAB19128.1; -.
EMBL; M38187; AAA48731.1; -.
ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
   +
                                                                                                                                                                                                                                                                 ---
```

```
STRAIN=C5;
MEDLINE=20063481; PubMed=10594704;
                                                                                                                                                                                                                                                                                                                                                                                                                                             1340 FGPSGEGSASGDLSGGAS 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last seq
16-OCT-2001 (Rel. 40, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Proteobacteria; Gamm
Xanthomonadaceae; Xanthomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                    722 IGTNGRGIVYGDIGGAPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=44291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QUIA XANCJ
Q9XD78;
                                                                                                                                                                           960
                                                                                                                                                                                                                                                                                    1197
                                                                                                                                                                                                                                                                                                                                       1250
                                                                                                                                                  456
                                                                                                                                                                                                     503
                                                                                                                                                                                                                                                                                                              604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
QUIA XANCJ
                                                                g
                                                                                           ò
                                                                                                                  셤
                                                                                                                                                 ò
                                                                                                                                                                     셤
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NTIMVAT------QISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSL 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | : |: :: | | : | |: | | : | | TSTNASIPEVSGEPESGEH---PTSGEPSASGAPDTSGEPTS-----VGFELSGEOSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        849 IGESG----LPSVDLQSSGFVPGESGLPSGDVSGLPSGIVDISGLPSAEEEVTVSVSRIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EVSGMPSGAESSGLHSGFSGEI-----SGT-ELISGLPSGEESGLASGFP--TISLVDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   202 LGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIAT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----SNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TDSGATWSQMT-NFPDVGTYIANPT-----DTTGYQSDIQGVVWVAFDKS--SSS
                                      PROSITE; PS01241; LINK, 4.
Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
Alternative splicing; Repeat; Immunoglobulin domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                 LINK 2.
LINK 3.
LINK 4.
19 X 20 AA TANDEM-REPEAT.
EGF-LIKE.
C-TYPE LECTIN.
                                                                                            AGGRECAN CORE PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Missing (in isoform /FTId=VSP 003073.
E -> D (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC.
                                                                                                        G-LIKE V-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY
SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY
                                                                                                                                                                                                                                                                                                                                                 BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
BY SIMILAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                              G1-A.
G1-B'.
G1-B'.
G2-B'.
KS.
CS-1.
CS-2.
G3.
 EGF_3; 1.
EGF_CA; 1.
IG_LIKE; 1.
LINK; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.6%
Query Match
Best Local Similarity 21.1%
Matches 143; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1870
1870
1890
1908
2017
2009
2067
                                                                                                                                                                                                                                                                                                           1264
1742
2109
129
242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             362
                                                                                                                                                                        1742
1892
2019
2081
137
243
PS50026;
PS01187;
PS50835;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1925
1993
2024
2053
76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      96/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                   ISULFID
                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                      DISULPID
                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                 DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                           DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
               PROSITE;
                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                SIGNAL
                                                                                                                     DOMAIN
                                                                                                                                    DOMAIN
                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                           DOMAIN
 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EESGLISGFPTVSLVDTTLVEVVTQTSVAQEVGEGPSGMIEISG---FLSGDRGVS---- 1249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |----GEGSGAVQSSGLPSGTGD-FSGEPSGIPYFSGDISGATDLSGQPSAVT---DISG 1300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GETAVTFPYV 1339
                                                                                       403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               558 NSWAASQGVPANAQIRSDRV-------NPKTFYALSNGTFYRSTDGGVTFQPV 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSA 663
-----EWDSSGARG- 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGSSYPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRVY 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lee Y.-A., Lo Y.-C., Yu P.-P.;
"A gene involved in quinate metabolism is specific to one DNA homology group of Kanthomonas campestris.";
J. Appl. Microbiol. 87:649-58(1999).
-!- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-dehydroquinate + reduced pyrroloquinoline-quinone.
-!- COPACTOR: PQQ (By similarity).
-!- ATHWAY: Quinic acid catabolic pathway; first step. This pathway allows growth of bacteria with quinate by its conversion to protocatechuate and subsequent metabolism by the beta-ketoadipate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404 YATNDLTKWDSG-----GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFT---
                                                                                                                                                                                                                                                                                                                                                                                                                             --HADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR-----AGSFDPSSQPND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RHVAFSTDGGKNWFQGSEPGGVT----TGGTVAASADGSRFVWAPGDPGQPVVYAVGFG
                                                                                       RYVLDISAEPWLTFGVQPNPP-----VPSPKLGWMDEAMAIDPFNSDRMLYGTGATL
                                                                                                                                                                     ---LPSGAE---TSGEQSGVPELSGEHSGVP----GLSGEAFEVPELSGEH----SGVT-
                                                                                                                                                                                                                                                                                                                                      ----ELSGEHSGLPELSGEPFGVPELSGFPSGL--DISGEPSGAPEVSGPVDVSGLTSGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -i- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Xanthomonas campestris (pv. juglandis).
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Probable quinate dehydrogenase (Pyrroloquinoline-quinone)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           790 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1301 EDSGLPEVTLVTSDLVEVVTRPTVSQELG---
```

Thu May

```
LSPEKGFVLG 324
                                                              Y48L SYNY3
P73069;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  522
                     RESULT 14
Y48L_SYNY3
                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                                                                                                                                                         50;
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                : | | | : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | | : | : | | : | : | | : | : | : | | : | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -FGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATW-----TRIWDWTSYPNRS 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          738
                                                                                                                                                                                                                                                                                                                                                                                                                                                60 WNNWGY--NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFK 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----IANPTDTTG-----YQSDIQG-VVWV-----AFDKSSSSLGQASKT 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      258 TSNTGGPYD-----GSS-----GDVW-KFSVTSGTWTRISPV--PSTD-----TANDY 297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLDPGNPHDRQAPAAGSSYVRSTPNVWAPMSYDAAMNTVFLPLGGPSTDLYGAERTALDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SDWWGATAMDQMLCRIAFKQMRYEGLYTAP----GTDVSL-----SF-PGSL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         627 EAVNTGMGAVPLKGTPYAVN-KNRFLSALGIPCQAPPYGTLSAIDLKTRSIAWQVPVGTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ODTGPFGIKMHLPIPIGMPTLGGTLSTQGGLVF--IAGTQ-DYYLRA----FDSATGKEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      209 IFVGVA-DP-----NNPVFWSRD--GGATWQAVPGAPT---GFIPHKGVFDPVNHVLYIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         351 LRYVLDISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADV---TAVPST----IFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GGMNW-----GGLSTDPVHDVVFANDMRLGLWVQMIPADTRKAEAAGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WSAYGNTDGGSRFAALDQINRSNGRPAAG---SPGPTTPGEIANSDGNGAEDQLTPLQ--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----https://www.common.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TKWDSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLG--GFTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------PGQPVVYAVGFGNSWAASQGVPANAQIRS--DRVNPKT---FYALSNGTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YRSTDGGVTFQ-PVAAGLPSSGAV----GVMFHAVPGKEGDLWLAASSGLYHSTNGGSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  118 LGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               501 NDRHVAFSTDGGKNWFQGSEPGGVTTGGT--VAASADGSRFVWA---PGD-----
                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                       Indels 275;
                                                                                                                                                                                                                                                                                                                                                                Length 790;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WSAITGVSSAVN-VGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WKGRLPVGSQGGPITYVSHKTGKQY----VVISAG----GARQSPDRG 778
                                                          EMBL; AF109471; AAD38453.1; -.
InterPro; IPR001479; Bac_PQQ.
InterPro; IRR002372; Bac_PQQ_repeat.
PF01011; PQQ; 2.
SMART; SM00564; PQQ; 2.
PROSITE; PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
PROSITE; PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
OXIGOTEQUECE PQQ; Quinate_metabolism; Transmembrane.
TRANSMEM 48 68 POTENTIAL.
TRANSMEM 48 68 POTENTIAL.
                                                                                                                                                                                                                                                                                                                     B75F29B52A49FE6F CRC64;
                                                                                                                                                                                                                                                                                                                                                            3.6%; Score 145.5; DB 1;
21.6%; Pred. No. 0.36;
                                                                                                                                                                                                                                                                                                                                                                                                     87; Mismatches 239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----DLPMOPSLIDFPNÓDGSHTPAVVIG-----
                                                                                                                                                                                                                                                                            POTENTIAL.
                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                   82896 MW;
                                                                                                                                                                                                                                                                                                                                                                                                       Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RYGASVLALD-
                                                                                                                                                                                                                                                                                                                     790 AA;
                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 546
                                                                                                                                                                                                                                                                              TRANSMEM
                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                 TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                      Local
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch). 18; 409 174 GAVGVARTIQRSTDG-RYVAVSARGNFYSTWAPGQTEWTPHNRNSSRRLQTWGYGKDGQL 232 ----GCDTW-----EQKLIDLGEEKASFSAVSFSGNEGW----ITGKPSIL--- 115 521 614 GVMFHAVPGKEGDLWLAASSG-LYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF 672 TKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTS 469 116 -LHTTDGGQTWARIPLSEKLPGAPYSIIALGPQTAEMITDLGAIYKTTNGGKNW-KALVE 173 GGVTTGGTVAASADGSRFV------WAPGDP------GQPVVYAVGFGNS--- 559 WAASQG-----VPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAV 613 ------SWGLL 267 77 ŝ PVFTTGTSVDYAEL-----NPSIIVRAGSFDPSSQPNDRHVAFSTDGGXNWFQGSEP SLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDL SCSHVPDLAFNPWQEIALETDST------FADIAFTEDP--NHGWLVGTKETIFETTD-Gaps SEQUENCE FROM N.A.
MEDLINE=97061201; PubMed=8905231;
KADRENTRE=97061201; PubMed=8905231;
KADREKO T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
Hosouchi T., Mateuno A., Muraki A., Nakazaki N., Naruo K., Okumura &
Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S., "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."; 3.6%; Score 145; DB 1; Length 342; 22.2%; Pred. No. 0.14; .ive 48; Mismatches 130; Indels 110; PIR; S75177; S75177. Interpro; IPR002860; GH BNR. PYam; PF02012; BNR; 4. BYpothetical proteome. SEQUENCE 342 AA; 37291 MW; 9AB2A72B82DA046A CRC64; Synechocystis sp. (strain PCC 6803). Bacteria; Cyanobacteria; Chroococcales; Synechocystis. NCBL_TaxID=1148; 15-JUL-1999 (Rel. 38, Created) 15-JUL-1999 (Rel. 38, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Res. 3:109-136(1996). SIMILARITY: Belongs to the ycf48 family. Ź EMBL; D90903; BAA17091.1; -. 82; Conservative STANDARD; 673 VVGTIGGVTG 682 16-OCT-2001 (Rel. 40 Ycf48-like protein. PIR; S75177; S75177 Query Match Best Local Similarity Matches 82; Conserv

```
ద
                           ò
                                                셤
                                                                                 ઠે
                                                                                                       g
                                                                                                                                    ò
                                                                                                                                                            g
                                                                                                                                                                                        ò
                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97 GINVVRVPVSTQLLLEWKAGTFLKPNVNTYANPELEGKNSLQIFEYWLTLCQKYGIKVFL 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
                                                                                                                                                                                                  MEDLINE=93209933; PubMed=8458833; Miller R.C. Jr., Warren R.A.; Meinke A., Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.A.; "Cellulose-binding polypeptides from Cellulomonas fimi: endoglucanase D (CenD) a family A beta-1,4-glucanase."; J. Bacteriol. 175:1910-1918 (1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DSTGKEAILSGVNWFGFN-----ASERVFHGL-----WS--GNITQITQQMAQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl hydrolases).
                                                                                                                                                                                                                                                                                     linkages in cellulose, lichenin and cereal beta-D-glucans. PATHWAY: Cellulose degradation.
SIMILARITY: Contains 2 fibronectin type III domains.
SIMILARITY: Contains 1 bacterial-type cellulose-binding (CBD)
                                                     01-OCT-1996 (Rel. 34, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Endoglucanase D precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Score 145; DB 1; Length 747;
; Pred. No. 0.36;
69; Mismatches 234; Indels 252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PS00659; GLYCOSYL HYDROL F5; 1.
PS00659; GLYCOSYL HYDROL F5; 1.
1 39 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DITIGYOSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGAT
                                                                                                                                    Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Micrococcineae; Cellulomonadaceae; Cellulomonas.
NCBI_TaxID=1708;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----WQA----VPGAPTGFIP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
PROTON DONGN (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
BD15473C9D8B42BD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ENDOGLUCANASE D.
                747 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001919; Bac_celose-bind. InterPro; IPR008965; Cellul bind. InterPro; IPR008957; FN III-like. InterPro; IPR003961; FN III. InterPro; IPR001547; Glyco_hydro_5. Pfam; PF00553; CBM 2; 1.
                                         (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78936 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.6%;
20.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; L02544; AAA23089.1; -. HSSP; P07986; 1EXG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam, PF00041; fn3; 2.
SMART, SM00637; CBD_II; 1.
SMART, SM00060; FN3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 141; Conservative
               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      546
208
349
747 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                        Cellulomonas fimi
                                       01-OCT-1996
01-OCT-1996
10-OCT-2003
                                                                                                Cellulase).
            GUND CELFI
P50400;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
GUND CELF1
                                                                                                                                                                                                                                                                                                                                                     +
                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
157 DVHSAEADNSGHVYNMWWKGDITTEDVYEG----WEWAAT--RWKDDDTIVGADIKNEP 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210 HGTQGSTERAKWDGTTDKDNFKHFAETASKKILAINPNWLVFVEGVEIYPKPGVPWTSTG 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   567 PA-NAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVAAGLP-----SSGAVGVMFH 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     382 VLNP-NSGDTGGL--LLDDWKTWDEVKYSTMLEPTLWKHGGKYVGLDHQVPLGGVGSTTG 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---GKNWFQGSEPGGVTTGG--TVAASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGV 566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   356 DISAEPWLTFGVQPNPPVPSPKLG------WMDEAMAIDPFNSDRMLYGTGATLY 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               328 DVWGPNWLFIHDEDIAPLLIGEWGGRLGODPRODKWM---AALRDLVAERRL---SQTFW 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                405 ATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISP---PSGAPLIS----ALGDLGGFTH 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          537 TARTAAGGGDVTAPSVPTGLTAGT------PTATSVPLTWTASTDTGGSGVTGY 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             585 EV--YRGSTLVARPTGTSHTVTGLSAATAYTFTVRAVDAAGNVS-AASAPVGVTTAPDPT 641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              484 GVAGYEVYRGTTLVGTTTATSYTVTGLAADSAYTFS-----VRAKDGAGNTSAASAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | : | | : | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                619 AVPGKEGDLWLAASSGLYHSTNGGSSWSAIT----GVSSAVNVGFGKSAP-----
--HKGVFDPVNHVL---YIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   457 ADVTAV----PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDG-
                                                                                                                                                                                                                                                                                                                                                                                                             ---HPNTIMVATQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              665 -GSSYPAVFVVGTIGGVTGAYRSDDCGTT----WVL 695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                642 TGSCAVTYTANGWSGGFTAAVTLINTGTTALSGWTL 677
                                                                                                                                                                                                                                                                                                                                                                                                  FGYSGL-----TIDRO---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                320
```

Search completed: May 11, 2004, 12:07:27 Job time: 25.1745 secs

```
GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

May 11, 2004, 11:45:55 ; Search time 86.3468 Seconds (without alignments) 2704.020 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-917-376-3 4036 1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

1017041 seqs, 315518202 residues Searched: 1017041 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		Description	Q82m04 streptomyce	086727 streptomyce	Q9aqh0 caldicellul	Q97kk0 clostridium	Q8tfpl aspergillus	Q9p4t8 agaricus bi	O74170 aspergillus	Q729m8 trichoderma	Q82k30 streptomyce	Q8plm5 xanthomonas	Q9wyel thermotoga	Q8p9u5 xanthomonas	Q8j0d2 geotrichum	Q8j1h7 agaricus bi	Q7x2u2 uncultured	Q9i2m3 pseudomonas
SUMMARIES		ΩI	Q82M04	086727	Q9AQH0	Q97KK0	QSTFP1	Q9P4T8	074170	Q7Z9M8	Q82K30	OBPLMS	Q9WYE1	Q8P9U5	Q8J0D2	Q8J1H7	Q7X2U2	Q912M3
		DB	16	16	~	16	n	m	m	ო	16	16	16	16	ო	m	7	16
		Match Length DB	882	890	966	839	857	806	856	838	739	751	707	751	812	196	1064	2468
	ا ا ا	Match	60.2	58.7	51.1	49.8	43.7	43.4	41.6	39.1	34.8	28.8	28.8	28.4	27.7	10.1	9.1	5.9
		Score	2429.5	2371	2062.5	2009	1765.5	1753.5	1680	1579	1405	1162.5	1160.5	1145	1117	409	368.5	238
	1.000	No.		7	n	4	5	9	7	80	σ	10	11	12	13	14	15	16

Q91948 Q98RG2 Q98RG2 Q7WNS4 Q92UU8 Q92UU8 Q91J59 P73139 P74440 Q971P0 Q97UP0 Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q98CG Q	091948 pseudomonas 088rg2 pseudomonas	Q7wbn0 bordetella Q92uu8 rhizobium m	Q7wn54 bordetella	Q8tjs9 methanosarc	P73139 synechocyst	P74440 synechocyst	Q9n1p0 bos taurus	Q9z4il bacillus sp	Q9hlq9 thermoplasm	Q88gz6 pseudomonas	Q8tjs8 methanosarc	Q9rk65 streptomyce	Q8xw00 ralstonia s	н	Q7uf08 rhodopirell	Q8fus1 brucella su	Q8brf6 mus musculu	Q80un3 mus musculu		O53393 mycobacteri		Q8ti55 methanosarc	Q8yyq5 anabaena sp				Q7wpl2 bordetella
	948 8RG2	WBN0 2008	WN54	TJS9	3139	4440	1P0	4I1	HLQ9	8GZ6	TJS8	RK65	XWOO	319	UF08	FUS1	BRF6	0UN3	OWRO	3393	8E20	TISS	YYQS	WIMI	UWZ8	YR62	WPL2
	16	16	16	17	16	16	9	~	17	16	17	16	16	σ	16	16	1	11	11	16	•	٠.	٠.	_	-	Н	16
	1800	2215	3346	2016	3972	4199	818	997	966	802	2523	848	1070	1904	2079	3420	999	928	1462	2523	3930	1271	993	344	7716	908	344
	ν. ω. ω.	5.0	4.9	4.9	4.7	4.6	4.6	4.6	4.6	4.6	4.5	4.5	4.5	4.5	4.4	4.4	4.4	4.4	4.4	4.3	4.3	4.3	4.3	4.2	4.2	4.2	4.2
18800 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	214	202.5	198.5	197	189	186.5	185.5	185.5	184.5	184	183	182.5	182	182	179	178.5	177	177	177	175.5	174.5	174	172	171.5	171.5	171	170.5
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	17	20	21	22	23	24	52	56	27	58	53	30	31	32	33	34	32	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

IEA.

ineae; Streptomycetaceae; Streptomyces.

Harris

FROM N.A.

```
Redenbach M., Kieger H.M., Denapaite D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.;
Kinashi H., Hopwood D.A.;
"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3 (2) chromosome.";
Mol. Microbiol. 21:77-96 (1996).
                                                                                                                                                                                                                                                                                                                                                                                                  Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S., Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Complete genome sequence of the model actinomycete Streptomyces
   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                        Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                      Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN-A3(2) / M145;
MEDLINE-21996410; PubMed=12000953;
                                                                                                                                                                                                                                STRAIN=A3(2);
MEDLINE=97000351; PubMed=8843436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         coelicolor A3(2).";
Nature 417:141-147(2002).
EMBL; AL999128; CAA20642.1; -.
PIR; T35237; T35237.
                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                              SEQUENCE FROM N.A.
                            NCBI_TaxID=1902;
                                                                            STRAIN=A3(2);
                                                                                                                                                          STRAIN=A3(2);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hopwood D.A.;
                                                                                             Seeger K.J.,
                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                          œ
                                                                                                                                                                                                                   125
                                                                                                                                                                                                                                                 158
                                                                                                                                                                                                                                                                               185
                                                                                                                                                                                                                                                                                                            218
                                                                                                                                                                                                                                                                                                                                          245
                                                                                                                                                                                                                                                                                                                                                                        278
                                                                                                                                                                                                                                                                                                                                                                                                   246 VFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTI 305
                                                                                                                                                                                                                                                                                                                                                                                                                                   335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                306 DROHPNTIMVATOISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     746
                                                                                                                                                       65
                                                                                                                                                                                     98
                                                                                                                                                                                                                                                                                                                                                         219 DNQGIVWVTFDESTGSPGSSTRTVYVGVADKDNSVYRSTDAGATWSRLAGQPTGHLAHKG
                                                                                                                                                                                                                                                                                                                                                                                                                       336 DRQHPGTVMATAYSSWWPDTQLFRSTDSGGTWTKAWDYTSYPSRSNRFTMDVSSSPWLTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GANPAPPEQTPKLGMMTESLEIDPFDSARMMYGTGATVYGTDNLTNWDSGSQFTIKPMAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          516 PGTVVRVGNLD--SGP---HVAFSTDNGANWFAGADPSGVSGGGTVAAASDGSRFVWSPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTG--VQYTTGFGTSWSASAGLPAGAIVESDRVDPKTFYGFKSGRFYVSSDGGATFTASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDCAILRSSDQGATWQITPLPFKLGGNMPGR
                                                                                                                                                                                                                                   TGVVSLASDSVDPNKVYAAVGTYTNSWDPGNGAVLRSGDRGASWQKTDLPFKLGGNMPGR
                                                                                                                                                                                                                                                                            GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                               GMGERLAVDPNRNSVLYLGAPSGKGLWRSTDSGASWSQVTDFPNVGTYVQDATDTSGYAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGOIHIAPMVK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              545 DPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           605 A-GLPSSGAVGVMFHAVPGKEGDLWLAASS----GLYHSTNGGSSWSAITGVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR
                                                                                                                                                     6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                  39 YSWKNARVDGGGFVPGIVFNRSEKNLAYARTDIGGAYRWAESSKTWTPLLDSVGWSDWGH
                                                                                                                                                                                                                                                                                                                                         DIOGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLEETAVNDLISPPS-GAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELN
                                                                                                                          Gaps
                                                                                                                          19;
                                                                                          DB 16; Length 882;
                                                                                                          Best Local Similarity 59.0%; Pred. No. 1.6e-126; Matches 438; Conservative 111; Mismatches 174; Indels
                                                            882 AA; 92152 MW; A3C8E753CE69F13D CRC64;
                                                                                          Score 2429.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |||: ||||||||||
RVYVSTNGRGIVYGDTAGSSDG 768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  719 RVYIGTNGRGIVYGDIGGAPSG 740
                                                                                          60.2%;
Pfam; PF02012; BNR; 10.
Pfam; PF00553; CBM 2; 1.
SMART; SM00637; CBD_II; 1.
                                            Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    485
                                                          SEQUENCE
                                                                                                                                                                                                                                              66
                                                                                                                                                                                                                                                                            126
                                                                                                                                                                                                                                                                                                          159
                                                                                                                                                                                                                                                                                                                                         186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               747
                                                                                          Query Match
S KW DR
                                                                                                                                                                                                                                         Вb
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                d
                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ⋩
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   220
                                                                                                                                                                                                                                                                                                                                                                                                                          125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     245 GVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLT 304
                                                                                                                                                                                                                                                                                                                                          65
                                                                                                                                                                                                                                                                                                                                                                        41 YTWKNARIDGGGFVPGIVFNRTEKDLAYARTDIGGAYRWQEESHTWTPLLDHVGWDDWGH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161 GMGERLAVDPHDNDVLYLGAPSGHGLWRSTDAGVTWSEVTAFPNPGNYAQDPNDTSGYAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                101 TGVVALASDAVDPDRVYAAVGTYTNDWDPTNGAVLRSADRGASWEKADLPFKLGGNMPGR
                                                                                                                                                                                                                                                                                                                                       6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                          66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DIOGVVWVAFDKSS-SSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHK
                                                                                                                                                                                                                                                                                            24;
HSSP; P07986; 1EXG.
GO; GO:000453; F:hydrolase activity, hydrolyzing O-glycosyl .
GO; GO:00045975; F:carbohydrate metabolism; IEA.
InterPro; IPR001919; Bac celose-bind.
InterPro; IPR008965; Cellul_bind.
InterPro; IPR002860; GH_BNR.
PF02012; BNR; 10.
                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                              D2C58695A4B56E84 CRC64;
                                                                                                                                                                                                                                                     58.7%; Score 2371; DB 16; 57.7%; Pred. No. 2.8e-123;
                                                                                                                                                                                                                                                                                              Matches 431; Conservative 109; Mismatches 183;
                                                                                                                                                                                                              93252 MW;
                                                                                                                                             Pfam; PF00553; CBM 2; 1.
SMART; SM00637; CBD_II; 1.
                                                                                                                                                                                        Complete proteome. SEQUENCE 890 AA;
                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               186
                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

Created)
Last sequence update)
Last annotation update)

01-NOV-1998 (TrEMBLrel. 08, 01-NOV-1998 (TrEMBLrel. 08, 01-OCT-2003 (TrEMBLrel. 25, Putative secreted cellulase.

OS GE DIT OS OS

Streptomyces coelicolor,

SC06545 OR SC5C7

890 AA

PRT;

PRELIMINARY;

086727; 086727 RESULT 2 086727

62

241

274

214

121

94

450 480 539

599 628

420

361

```
95 WNLLGVESIATDEVDENXVILACQCTYTNSWTDMNGAILRSTDEGDTFEITPLFFXLGGN 154
                                                                                                                                                                                                                                                             FCKSAPGSSYPAVFVVGTIGGVTCAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTT 181
                                                                                                                                                                                                                                              PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYS 301
                                                                                                                                                                                                        35 QPYVWKNVKIEGGGGPITGIVENPKEKNLVYVRTDIGGAYRSTDGGNTWTQLMDWVSFDD
                                                                                                                                                   182 GYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI
                                                                                                                                                                                                                                                                                                                                                                       421 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                       KVKAIGIEETSVQALISPPVGPHLFSALGDIAGFRHEDLEKAPNWTYVQPNMGTTTDIDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F-QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVG
                              4 OPYTWSNVAI -GGGGFVDGIVFNEGAPGILYVRIDIGGMYRWDAANGRWIPLLDWVGWNN
                                                                                63 WGYNGVVSIAADPINTNKV-WAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGN
                                                                                                                                                                                                                                                                                                   GLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEP
                                                                                                                                                                                                                                                                                                                                                         WLTFG-VQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQG-SEPGGVTTGGTVAASADGSRF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clostridium acetobutylicum.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
    15;
    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
, several ASP-boxes and
    96; Mismatches 238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
MEDLINE=21359325; PubMed=11466286;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ||:: ||| || :|:|
RVFVATNGLGIKWGEI 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RVYIGTNGRGIVYGDI 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-0CT-2001 (TYEMBLrel. 18, C
01-0CT-2001 (TYEMBLrel. 18, L
01-0CT-2003 (TYEMBLrel. 25, L
Probably secreted sialidase,
    387; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=1488;
                                                                                                                                        122
                                                                                                                                                                                                                      215
                                                                                                                                                                                                                                                242
                                                                                                                                                                                                                                                                                                     302
                                                                                                                                                                                                                                                                                                                                                         362
                                                                                                                                                                                                                                                                                                                                                                                                                                       451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629
                                                                                                                                                                                                                                                                           275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q97KK0
Q97KK0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               domain
    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q97KK0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                  ò
                                                                                                        셤
                                                                                                                                       ઠે
                                                                                                                                                             셤
                                                                                                                                                                                          ò
                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                                                                                              --EGAGVQYTTGFGTSWQASTGLPAGAIVESDRVNPATFYGFKSGRFYVSTDGGATFTAS 630
                                                                                                                                                                                                                                                                                                                  657
                                                                                                                                                                                                                                                                                                                                 A BENGALINE=20171169; PubMed=10706665; Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., A Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., A Bergquist P.L.; Thermophile Caldicallucational glycosyl hydrolases from the extreme thermophile Caldicallucations isolate Tok7B.1."; Curr. Microbiol. 40:333-340(2000). EMBL; AF078038; AAK06388.1; Curr. Microbiol. 40:333-340(2000). R RISSP; Q06881; 1NBC. GO; GO:0016787; F:hydrolase activity; IEA. GO; GO:0016787; F:hydrolase activity; IEA. RISSP; Q06881; INBC. R GO; GO:001956; CBD 3. R InterPro; IPR001956; CBD 3. R InterPro; IPR001956; CelTul_bind. R InterPro; IPR002860; GH_BNR. PFam; PF00042; CBM 3; 1.
GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANL 717
                                                                                              FGVOPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW-DSGGOIHIAPM 423
                                                                                                                                                     VKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAEL 483
                                                                                                                                                                                                                                                              544 GDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV 603
                                           305 IDROHPNTIMVATOISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLT 364
                                                                                                                                                                                                        484 NPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAP
                                                                                                                                                                                                                       518 KPDVVVRAGNLD--SGP---HIAFSTDNGANWFGGTDPSGVSGGGTVAAGADGSRFVWSP
                                                                                                                                                                                                                                                                                                                  AA-GLPSSGAVGVMFHAVPGKEGDLWLAASS----GLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Glycosyl hydroclase 5 (Fragment).
Caldicellulosiruptor sp. Tok/B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               996
108275 MW; 3C72B6ED22F3C614 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 2062.5; DB 2
Pred. No. 4.1e-106;
                                                                                                                                                                                                                                                                                                                                                                                                                                          749 GRVYVATNGRGVIYGDTSDTGGGTDPG 775
                                                                                                                                                                                                                                                                                                                                                                                                                            RRVYIGTNGRGIVYGDI - - - - GGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      51.1%;
52.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            996 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. STRAIN=TOk7B.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase.
NON TER
SEQUENCE
                                                                                              365
                                                                                                                                                     424
                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                                                          718
                                                                                                                         398
                                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                                                       573
                                                                                                                                                                                                                                                                                                                                            631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                              셤
                                                                                      à
                                                                                                                   셤
                                                                                                                                                   ò
                                                                                                                                                                       셤
                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                         8
```

```
SEQUENCE FROM N.A.
                                                                                                                                                    Aspergillus niger
                                                                                                                                                                                     NCBI_TaxID=5061;
                                                                                                                              Endoglucanase C.
                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S:
Matches 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     314
                                                                      Q8TFP1
                                                 S
                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                       12;
                                                                                                                                                                                                                                                                                                                       WGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNM 122
                                                                                                                                                                                                                                                                                                                                               154
                                                                                                                                                                                                                                                                                                                                                                                          212
                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                      272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDTTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                            PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYS 301
Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Tatusov W.E., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Batusov W.E., Shathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.M., Koonin E.V., Smith D.R., "Genome sequence and comparative analysis of the solvent-producing IT bacterium Clostridium acetobutylicum.";
J. Bacteriol. 183:4823-4838(2001).
BMBL, AE007608; AAX78895-1; -.
BMBL, AE007608; P. Fcalcium ion binding; IEA.
CO, GO:000520; P:calcium ion binding; IEA.
BIR: D97013; D97013.
BY GO, GO:000272; P:polysaccharide catabolism; IEA.
InterPro; IPR002048; EF-hand.
InterPro; IPR002048; EF-hand.
BR InterPro; IPR002048; EF-hand.
BR Pfam; PF00404; Dockerin 1; 2.
BR Pfam; PF00404; Bockerin 1; 2.
BR Pfam; PF00404; Bockerin 1; 2.
BR PSGYIES; PS00018; EF-HAND; 2.
                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                                  94
                                                                                                                                                                                                                                                                                         SQGYKWDNAKIGAGGYVPAVIFNKTEKDLIYARTDMGGAYRWDKANNKWIPITD--GFSD
                                                                                                                                                                                                                                                                                                                                   3 TQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNN
                                                                                                                                                                                                                                                                                                                                                                                                               YQSDIQGVVWVAFDKSSSSLGQASKTIFVGVAD-PNNPVFWSRDGGATWQAVPGAPTGFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAASADGSRF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 509 AESVPNFVVRVGNVDTSKNQQDKDCGISYDGGKNWFSAGSNISGVYKAGTVAAGADAKTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F-OPVAAGLPSSGAVGVMFHAVPGKEGDLWLA-ASSGLYHSTNGGSSWSAITGVSSAVNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FTQSSQTGLPTKGK-GI-FKTVIGHEGDIWIAGGKDGLWHSTDSGATFTKVSGVDASDTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  658 GFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WLTFG-VQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGOIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APMVKGLEETAVNDL1SPPSGAPL1SALGDLGGFTHADVTAVPST1FTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        540 VWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVT
                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                       16;
                                                                                                                                                                                                                                 DB 16; Length 839;
                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                         839 AA; 90824 MW; 6C99A041CA9CF984 CRC64;
                                                                                                                                                                                                                                         Best Local Similarity 50.9%; Pred. No. 3e-103; Matches 377; Conservative 107; Mismatches 240;
                                                                                                                                                                                                                              49.8%; Score 2009; 50.9%; Pred. No. 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RRVYIGTNGRGIVYGDIGGA 737
                                                                                                                                                                                               Complete proteome
SEQUENCE 839 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            718
                                                                                                                                                                                                                                                                                                                       63
                                                                                                                                                                                                                                                                                                                                                                                         155
                                                                                                                                                                                                                                                                                                                                                                                                               183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     009
                                                                                                                                                                                                                                                                                                                                                                   123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           628
                                                                                                                                                                                                                                                                                                37
                                                                                                                                                                                                                                                                                                                                             95
                                                                                                                                                                                                                                                                                                                                                                                                                                                            242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      449
                                                                                                                                                                                                                               Match
                                                                                                                                                                                                                                 Query
   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    374 NAPWLQDDTSTDEFPV---RVGMWPEALAIDPFDSDHWLYGTGETIYGGHDLQNWDSEHN 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIANPTD 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYF 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDIS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           359 AEPWLIFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQ 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTS 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20 AASQAYTWKNVVTGGGGGFTPGIVFNPSAKGVAYARTDIGGAYRLN-SDDTWTPLMDWAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 ATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              T--YTSDPVGIAWVTFDSTSGSSGSPTPRIFVGVVDTGESVFVSEDAGETWTWVSGEPMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "EglC, a New Endoglucanase from Aspergillus niger with Major Activity
towards Xyloglucan.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29; Gaps
                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=21914059; PubMed=11916668;
Hasper A.A., Dekkers E., van Mil M., van de Vondervoort P.J.I.,
de Graaff L.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0-glycosyl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90459 MW; 482253ECD919C511 CRC64;
                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Appl. Environ. Microbiol. 68:1556-1560(2002).

EMBL, AY040839; AAK77227.1; -...

GO; GO:0005576; C:extracellular; IEA.

GO; GO:0004553; F:hydrolase activity, hydrolyzing

GO; GO:0005975; P:carbohydrate metabolism; IEA.

InterPro; IPR00254; CBD fungal.

InterPro; IPR002860; GH_BNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 / Match 43.7%; Score 1765.5; DB Local Similarity 47.1%; Pred. No. 9.6e-90; Nes 351; Conservative 107; Mismatches 258
                                                                                                                                                                   Ź
                                                                                                                                                                 857
                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF02012; BNR; 7.
Pfam; PF00734; CBM 1; 1.
ProDom; PD001821; CBD fungal; 1.
PRART; SM00236; CBD; 1.
PROSITE; PS00562; CBD FUNGAL; 1.
SEQUENCE 857 AA; 90459 MW; 4!
PRELIMINARY;
                                                                                                                                                                                                                    01-JUN-2002 (TrEMBLrel.
01-JUN-2002 (TrEMBLrel.
01-JUN-2003 (TrEMBLrel.
```

요

ઠ

8 & 8

ò

8 & 8

ò

```
589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            655
                                                 261
                                                                                                                                                                                                                            415
                                                                                                                                                                                                                                                                                                                                                   434
                                                                                                                                                                                                                                                                                                                                                                                                                                 551
                                                                                                                                                                                                     GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDIS 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IEA
                                                                                                                                                                                                                                                                                                                              TTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADP-NNPVFWSRDGGATWQAVPGAPT
                         GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYF
                                                                                                                            262 SFLPHKGVLSPSERALYITYSDGGGPYDGTSGAVXKYSIDTGAWTDITPVSGSDL---FF
                                                                                                                                                                                                                                                                                                  359 AEPWLTFGVQPNPPVPSP---KLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSG
                                                                                                                                                                                                                                                                                                                                                                                                    416 GQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         476 TSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSRFVWAPPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        552 ADIILWRINGNG--------WWSRNQATFNIFYGASGKTFYVSTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            596 GGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              656 NVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYG-NWGQAITGDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          648 GVALGAPRSTGGYPAVFAAANYGGEVAYFRSDDRGVNWVKINDAAHGFGAASANCMAADP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Aspergillus aculeatus.
Eukaryota, Pungi, Ascomycota, Pezizomycotina, Eurotiomycetes;
Eurotiales, Trichocomaceae, mitosporic Trichocomaceae, Aspergillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Arain M., Takada G., Kawaguchi T., Sumitani J.;
"Avicelase III from Aspergillus aculeatus.";
"Lu Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
BRBL; ABO1551; BAA29031.1; -.
BRBL; ABO1551; BAA29031.1; -.
BRSP; P00725; 2CBH.
RSSP; P00725; 2CBH.
RGO; GO:0005576; C:extracellular; IEA.
GO; GO:0005576; C:extracellular; IEA.
GO; GO:0005576; P:hydrolase activity, hydrolyzing O-glycosyl ..
RGO; GO:0005575; P:hydrolase activity, hydrolyzing O-glycosyl ..
RGO; GO:0005576; P:hydrolase activity, hydrolyzing O-glycosyl ..
RGO; GO:0005576; C:extracellular; IEA.
RGO; GO:0005576; C:extracellular; IEA.
RGO; GO:0005576; C:extracellular; IEA.
RGO; GO:0005576; C:extracellular; IEA.
RGO; GO:0005576; C:extracellular; IEA.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:0005576; C:extracellular; I.
RGO; GO:000576; C:extracellular; III.
RGO; GO:000576; C:extracellular; III.
RGO; GO:000576; C:extracellular; III.
RGO; GO:000576; C:extracellu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RVYGRVYIGTNGRGIFYGDVAGS 730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 715 ANLRRVYIGTNGRGIVYGDIGGA 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08,
08,
24,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1998 (TrEMBLrel.
01-NOV-1998 (TrEMBLrel.
01-JUN-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=5053;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avicelase III.
180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  708
                                                                                                  239
                                                                                                                                                                                                     299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         074170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVIIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                         셤
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLG 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S2 DSRMNYMGVDALATDPVQDNRLYLATGLYTNSWDPSNGHILISSDYGKTFDAAPLPFKLG 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              597
                                                                                                                                                        602 ATF-TVTTIGSSTTANAI-RAQPSLAGDVWVSTDTGLFHSTNYGKSFTQIGSGCTEGWS 659
                                                                                                                                                                                                                                                                                                                                                        660 FGFGKPSSDGDYPVLFGFFTVDGVTGLFKTEDQGVNWQIISDAEHGFGSASANVVNGDLQ 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTD 179
                                                                                                                                                                                                                         VTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSSAVN 656
                                                                                                                                                                                                                                                                                                                            657 VGPGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDOHQYGN-WGQAITGDHA 715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23 AASÓSYRWONVKIGGGGGFVPGIVFNPSBKGLAYARTDIGGAYKLN-ADDTWTPLLDPAD 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 ATTOPYTWSNVAI - GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                       538 RFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Morales-Almora P., Thurston C.F.;
"Molecular analysis of the callulolytic genes in Agaricus bisporus.";
Submitted (ANG-2000) to the EMBL/GenBank/DDBJ databases.
BRMBJ: AJ792929; GC02964.1;
HSSP; P00725; 2CBH.
                         VDYAELNPSIIVRAGSFDPSSOPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Agaricus bisporus (Common mushroom).
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Agaricaceae; Agaricus.
NCBI_TaxID=5341;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GO; GO:0005576; C:extracellular; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR000254, CBD fungal.
InterPro; IPR002860; GH_BNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 3; Length 806;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 43.4%; Score 1753.5; DB 3; Length Best Local Similarity 45.9%; Pred. No. 4.1e+89; Matches 341; Conservative 112; Mismatches 249; Indele
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7C67E382E62FF341 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15, Created)
15, Last sequence update)
24, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL. CEL6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        716 NLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00562; CBD FUNGAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84779 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF02012; BNR; 8.
Pfam; PF00734; CBM 1; 1.
SMART; SM00236; fCBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2000 (TrEMBLrel. 01-JUN-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23
806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24 8
806 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-OCT-2000
01-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=D649;
                         478
                                                                                                                                                                                                                         598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09P4T8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9P4T8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gnal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CEL6
```

PARTY KER TO DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND DE STAND

8 8 8 8 8

```
ø
                                                                                                                                                                                                                                                                                                                                                            80
                                                                                                                                                                                                                                                                                                                                                                                                                              137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     592
                                                                                                                                                                                                                                                                                          21
                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         662
                                                                                                                                                                                       Query Match
                                                                                                                                                                                                       Local
                                                                                                                                                                                                       Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
Q82K30
ID Q82
AC Q82
DT 01-
В
                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                   116
                                                                                                                                                                                     136
                                                                                                                                                                                                                       176
                                                                                                                                                                                                                                                         196
                                                                                                                                                                                                                                                                                          236
                                                                                                                                                                                                                                                                                                                            254
                                                                                                                                                                                                                                                                                                                                                            295
                                                                                                                                                                                                                                                                                                                                                                                            311
                                                                                                                                                                                                                                                                                                                                                                                                                                355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  717
                                                                                59
                                                                                                                 79
                                                                                                   SSST--YTSDPVGIAWVTFDSTSGSSGSATPRIFVGVADAGKSVFKSEDAGATWAWVSGE
                                                                                                                                                                                                                                                                                                                                                                                                                              DYFGYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GWSFGFGKASSTGSYVVIYGFFTIDGAAGLFKSEDAGTNWQVISDASHGFGS-GSANVVN
                                                                                                                                                   ---WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPP
                                                                                                                                                                                                                       KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN
                                                                                                                                                                                                                                                                                         PTDTTGYOSDIOGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGA
                                                                                                                                                                                                                                                                                                                                                            PT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KHNVIVKSLAVGIEEMAVLGLITPPGGPALLSAVGDDGGFYHSDLDAAPNOAYHTPTYGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TNGIDYAGNKPŚNIVRSGASD--DYPT---LALSSNFGSTWYADYAASTSTGTGAVALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT
                                                                                1 ATTQPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                                                                                                                                                                                                                                                                                              ---TSLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DISAEPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI -TGVSS
                                                 Gaps
                                                32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0729M8;
01-0CT-2003 (TrEMBLrel. 25, Created)
01-0CT-2003 (TrEMBLrel. 25, Last sequence update)
01-0CT-2003 (TrEMBLrel. 25, Last annotation update)
01-0CT-2003 (TrEMBLrel. 25, Last annotation update)
CEL74A.
Trichoderma reesei (Hypocrea jecorina).
Trichoderma reesei (Hypocrea jecorina).
Bukaryota; Fungi; Ascomyocta; Pezizomyoctina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
                                                                                                                                                                                                                                                                                                                                                                              Indels
                Length
            Score 1680; DB 3; 1
Pred. No. 5.2e-85;
3; Mismatches 258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GDLQTYGRVFRGHERPGHLLRQSQREPAG
                                              Matches 346; Conservative 113;
              41.6%;
                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE 1
                                                                                                                                                                                     80
                                                                                                                                                                                                                                                       137
                                                                                                                                                                                                                     117
                                                                                                                                                                                                                                                                                         177
                                                                                                                                                                                                                                                                                                                         197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  712
              Query Match
Best Local (
                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                          237
                                                                                                                                                                                                                                                                                                                                                                                            255
                                                                                                                                                                                                                                                                                                                                                                                                                              296
                                                                                                                                                                                                                                                                                                                                                                                                                                                              312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q729M8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
0729M8
0729M8
AC 0729
DT 01-0
DT 01-0
DT 01-0
DT 01-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
DT 02-0
                                                                                                          g
                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                ð
                                                                                                                                                ò
                                                                                                                                                                        g
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      313
                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                            136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISA 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        661
                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                   79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         197 NGYNSDKQGLMWVTFDSTSSTTGGATSRIFVGTADNITASVYVSTNAGSTWSAVPGQPGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHA----DVTAVPSTIFTSPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RHNVSIQSLADGIEEFSVQDLASAPGGSELLAAVGDDNGFTFASRNDLGTSPQTVWATPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGYOSDIOGVVWVAFDKSSSLGQASKT1FVGVADP-NNPVFWSRDGGATWQAVPGAPTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EPWLTFG-VQPNPPVPS----PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RSTDGGVTFQPVAAGLPSSGAVGVM--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        606 VSKDTGSSF---TRG-PKLGSAGTIRDIAAHPTTAGTLYVSTDVGIFRSTDSGTTFGQVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA--YRSDDCGTTWVLINDDQHQYGNW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TALTNIYQIALGVGS-GSNW-NLYAFGI--GPSGARLYASGDSGASWIDIQGSQGFGSID
                                                                                                                                                                                                                                                                                                                           YTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWV----GW
                                                                                                                                                                                                                                                                              Gaps
Foreman P.K., Brown D., Dankmeyer L., Dean R., Diener S., Dunn-Coleman N.S., Goedegebuur F., Houfek T.D., England G.J., Kelley A.S., Meerman H.J., Mitchell T., Mitchinson C., Olivares Teunissen P.J.M., Yao J., Marx M.;

"Transcriptional Regulation of Biomass-Degrading Enzymes in the Filamentous Fungus Trichoderma reesel.";

J. Biol. Chem. 0.0-0(2003).

EMBL; AY281371; AAP57752.1;

SEQUENCE 838 AA; 87132 MW; FBE3D08F2A74FA84 CRC64;
                                                                                                                                                                                                                                                                            44;
                                                                                                                                                                                                                                 Length 838;
                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              740
                                                                                                                                                                                                                             tch 39.1%; Score 1579; DB 3; 1 al Similarity 43.0%; Pred. No. 2e-79; 325; Conservative 123; Mismatches 264;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STKVAGSGSTAGQVYVGTNGRGVFYAQGTVGGGTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GQAITGDHANLRRVYIGTNGRGIVY -- GDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ŕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q82K30 PRELIMINARY;
Q82K30;
01-JUN-2003 (TrEMBLrel. 24,
```

581 583

DDT REPRESENT OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVICE OF SERVI

```
FTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTTGGT
                              | : | : | | : | | | : | | 475 FGSATGLAQAAARPAYVVRTGWGDHGNG-----AYSHDGGRTWAPPEAQPDIAKDAPGP
                                                                                                                            530 VAASADGSRFVWA----PGDPGQPVVYA----VGFGNSWAASQGVPANAQIRSDRVNPKT
                                                                                                                                                                                            529 IATSADGGTLLWSFVHWDG-----TTYAAHRSTDNGASWSEVSSFPKGATPVADPADPTR
                                                                                                                                                                                                                                                         FYA--LSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAAS-SGLYHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D4134D96285E599D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | ||:| |:|: || DAHQWGWIGEAVVGDPRIHGRVYLATNGRGIQYGE 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOHOYGNWGQAITGDHANLRRVYIGTNGRGIVYGD 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nature 417:459-463(2002).
EMBL; AE011809; AAM36634.1; -
InterPro; IPR002860; GH_BNR.
Pfam; PF02012; BNR; 7.
Complete proteome.
SEQUENCE 751 AA; 81259 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81259 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-OCT-2002 (TrEMBLrel. 22, 01-OCT-2002 (TrEMBLrel. 22, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=92829;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELA OR XAC1770.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             669
472
                                                                                                                                                                                                                                                            582
                                                                                                                                                                                                                                                                                                                                                                                             639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q8PLM5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q8PLM5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HIDDER REPAREMENT OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                            임
                                                                                                                                ઠે
                                                                                                                                                                               요
                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                             ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              102 LGVEAMAVDPTHPDRLYLAVGTYAQSWAGN-GAVLRSEDRGATWTRTDLTVKLGGNEDGR 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      186 DIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNP----VFWSRDGGATWQAVPGAPTGF 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----APRIRGLEESAVROLISPPVGEAHLISGLGDIGVMYHERLTASPSRGMATNPV 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPPKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SADTFAYGGVAVDARRPGTLVVSTNNRWADGDTVFRSTDGGRTWTSLKD-----AA 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VLDISAEPWLIFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWD 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SGGOIHIAPMVKGLEETAVNDLISPPSG-APLISALGDLGGFTHADVTAVPST-IFTSPV 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGOGVVF-----LVAAGRTVYAGWGDGDGTSGTANLYRTAD-GTTWGAVPGRPSGT 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 ---IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPV----PSTDT 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ANDYFGYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRY 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
Sakaki Y., Hattori M., Omura S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.";
Nat. Biotechnol. 21:526-531(2003).
EMBL; APOGS031; BAC70285.1; --
GO; GO:0016787; F:hydrolase activity; IEA.
InterPro: IPR003860; GH BNR.
InterPro: IPR003860; GH BNR.
Figan; PF02012; BNR; 9.
TIGREPMS; TIGR04409; TAT signal_seq; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 34.8%; Score 1405; DB 16; Length 739; Best Local Similarity 40.8%; Pred. No. 7.5e-70; Matches 308; Conservative 104; Mismatches 257; Indels 86
                                                                                                                                                                                                                                                                                                                  PERGINE AND MACC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=21477403; PubMed=11572948;
OWNERS: Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
"Genome sequence of an industrial microorganism Streptomyces avermitiiis: deducing the ability of producing secondary
                                                                                                                                                              Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase, Complete proteome.
SEQUENCE 739 AA; 78399 MW; 8E3B0F3D0F5B3D28 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
                              Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001)
   sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=22608306; PubMed=12692562;
01-JUN-2003 (TrEMBLrel. 24, Last 01-OCT-2003 (TrEMBLrel. 25, Last
                                     or-ocr-2003 (TrEMBLrel. 25,
Putative glycosyl hydrolase.
                                                                                                                                   Streptomyces avermitilis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                              NCBI_TaxID=33903;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
```

g

d

ò 셤 ò

8 셤 ð g ò g ò 셤

ò

ò

```
EQUENCE FROM N.A.

SEQUENCE FROM N.A.

RECUENCE TRAIN=306 / ATCC 13902 / XV 101;

RECUENCE STRAIN=306 / ATCC 13902 / XV 101;

RECUENCE STRAIN=306 / ATCC 13902 / XV 101;

RECUENCE STRAIN=306 / ATCC 13902 / R. Reinach F.C., Farah C.S., Furlan L.R.,

Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeda N.F.,

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,

RA Camarotte G., Cannavan A.M., Ferreira R.C.C., Ferro M.I.T.,

RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,

RATSUNGAMA A.M., Kishi L.T., Leite R.P., Lemos G.M., Lemos M.V.F.,

RATSUNGAMA A.M., Novo M.T.M., Madeira A.M.B.N., Martinez-Rossi N.M.,

Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,

RATINGA Gossi A., Sena J.A.D., Silva C., Oliveira V.R.,

Perreira H.A., Rossi A., Sena J.A.D., Silva C., Ge Souza R.F.,

RIMINGA Gos Santos M., Truffi D., Teai S.M., White F.F.,

Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,

RCOMPALISON Of the genomes of two Xanthomonas pathogens with differing M. Novelle S.M.,

Thost specificities.
                                                                                 698
                                                                                                                                 62 NWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGN 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   584 FYAYDFDNGTLYASTDSGRSFTARAGGLP-SGDSQFKLVAAPGRSGDLWLSAKWNGLYRS
                                                                         TNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 TTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Xanthomonas axonopodis (pv. citri).
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xanthomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tch 28.8%; Score 1162.5; DB 16; Length 751; al Similarity 35.7%; Pred. No. 2.1e-56; 268; Conservative 115; Mismatches 306; Indels 61;
```

œ

```
TIGR; TM0305;
                                                                                                                                                                                                                                                                                        99
                                                                                                                                                                                                                                                                                                                                                                          123
                                                                                                                                                                                                                                                                                                                                                                                                                   140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8P9US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12
S K B B B
                                                                                                                                                                                                          ઠ
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                     241
                                                                                                                                            279
                                                                                                                                                                                     294
                                                                                                                                                                                                                             330
                                                                                                                                                                                                                                                                      354
                                                                                                                                                                                                                                                                                                               382
                                                                                                                                                                                                                                                                                                                                                      414
                                                                                                                                                                                                                                                                                                                                                                                               430
                                                                                                                                                                                                                                                                                                                                                                                                                                      474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    720
                     181
                                           Melson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
"Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
Mature 399:323-329(1999).
BMBL, AE001112; AA015393.1; -.
                                                                                                                                                                                                                                                                    295 NDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYV
                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVAAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADGSRFVWAPGDPGQPVVYAVG-FGNSWAASQGVPANAQIRSDRVNPKTFYALS--NGTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 || |: || | | : : || | | : : || | : : : || | : : : || | : : || | : : || | : : : || | : : : || | : : : || | : : || | : : : || | | : : : || | : : : || | : : : || | : : : || | : : : || | : : : || | : : : : || | : : : : || | : : : : || | : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : : || : : : || : : : || : : : || : : : || : : : || : : : : || : : : : || : : : : || : : : : || : : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : : || : : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---AITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CIRTLSKSDEARSLGIGKALRAGAPPALYLAGRVAGVDGIFRSDDEGAHWRRINDDAHRF
                   MPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDTT
                                                                                                                                                                                  PHKGVFDPVNHVL----YIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTA
                                                                                                                                                                                                                                                                                                                                                    LDISAEPWLTFGVOPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                    GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YRSTDGGVTFQPVA--AGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                606 YESSDGAASFRDTGQQVGSPARDE----RARPQLRPDPWRAGVVYLASPTLGVMRWQNG
                                                                                                   GYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI
                                                                                                                                                                                                                                                                                            Thermotoga maritima.
Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga
NCBI_TaxID=2336;.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ā
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GK-PYSVTGDPRIAGRVYFATGGRGIFYGD 749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GNWGQAITGDHANLRRVYIGTNGRGIVYGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN=MSB8 / DSM 3109;
MEDLINE=99287316; PubMed=10360571;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Endoglucanase, putative.
                                                                                                                                                                                                                           ь.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         591
                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                  242
                                                                                                                                                                                                                           280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        534
                                                                                                                                         221
                                                                                                                                                                                                                                                                                                                                                                                             383
                                                                                                                                                                                                                                                                                                                                                                                                                                      415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         O9WYE1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9WYE1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                 g
                                                                                                                                                                                ઠે
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                 8
                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
```

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLP---FKLGGNM 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTG 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 183 YQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIP 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HKGVFDPVNHVLY1ATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSG 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPW 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAP 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  340 IK---DLNP------HMISD-VKIDPFDMNRAIFTTGYGVWVTYELKKSFEG----MGK 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          385 PVKWIFENRGLEETVVLQLVPPIGERPLLSAIADWGGFRHESLDTPPSSMY-KPLKWTSL 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGG--TVAAS 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      534 ADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPA---NAQIRSDRVNPKTFYAL--SNG 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               498 NDGKTLVWSPAN--HEVIVSSDKGKSWKKAISVPVPEFNYFPASDPVNPSKFYIFDWKNG 555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589 TFYRSTDGGVTFQPVAAGLPSSGAVGVMFHA----VPGKEGDLWLAAS-SGLYHSTNGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DFLISKDGGKSFWK-GAKLPSFDNWWVSLYSFPVLAPDREGDIWLALQWNGLYRSKDGGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MVK-----GLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                           6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                DB 16; Length 707;
                                                                                                                                                                                                                                                                                                                                             89;
                                                                                                                                                   8CD8743CBDA6A99F CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                      Query Match 28.8%; Score 1160.5; DB 16
Best Local Similarity 35.2%; Pred. No. 2.5e-56;
Matches 264; Conservative 123; Mismatches 275;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            704 GNWGQAITGDHANLRRVYIGTNGRGIVYGDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  G-WIHYMIGDMNEFGRIFLGTEGRGIIVGEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8P9U5;
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequen
01-UJV-2003 (TrEMBLrel. 24, Last annota
Cellulase.
XCC1752.
Xanthomonas campestris (pv. campestris)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                   79496 MW;
              GH_BNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
InterPro; IPR002860; C
Pfam; PF02012; BNR; 9.
Complete proteome.
SEQUENCE 707 AA; 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1 208 P9US 1D 208 P9US DT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 PDT 201 P
```

σ

710

652

604

```
TNGESIDAAGQAPQWVVRSGTV-RDRRNNEIRALYSRDGGKQWTAFASEPPAGQGAGSIA 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 NGVVSIAADPINTNKVWAAVGMYT-NSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87 MGTESIALDPNNPDRLYLAQGRYVGDEW----AAFYVSEDRGQSFTIYESPFPMGANDMG 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S-----VIFDP-----ERNGTIYASATAPQG-MYVTHDGGVSWEPVAGQPSSWLNRT 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Purification, Characterization, Cloning, and Expression of a Novel Xyloglucan-specific Glycosidase, Oligoxyloglucan Reducing End-specific Cellobiohydrolase.";
                                                                                                                                                                                                                                                547 IGADAAQVVWAPERGGN--WRTSDFGAQWQRVDGLPDTAVVWADRVDARRWYAVDVASGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQ-GSEPGGVTTGGTVA
                                                                                                                                                                           ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYA--LSNGT
                                                                                                                                                                                                                                                                                                                      -----PSSGAVGVMFHAVPGKEGDLWLAASS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                GLYHSTNGGSSWSAI TGVSSAVNVGFGKSAPGSSYPAVFVVGTI GGVTGAYRSDDCGTTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI---
                                                                                                                                                                                                                                                                                                                                                                                       605 LYESTDAARSFR--ATGVQVGSPARDERTRPQLRPDPWRAGVVYLASPGK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-JUM-2003 (TrEMBLrel. 24, Last annotation update)
01.JUM-2003 (TrEMBLrel. 24, Last annotation update)
01.jooxyloglucan reducing end-specific cellobiohydrolase.
Geotrichum sp. ML28.
Bukaryota, Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Dipodascaceae; mitosporic Dipodascaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27.7%; Score 1117; DB 3; Length 812; 35.0%; Pred. No. 7.7e-54; ive 124; Mismatches 279; Indels 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 812 AA; 87080 MW; F5C740F32D8B17E5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   694 VLINDDOHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      J. Biol. Chem. 277:48276-48281 (2002).

EMBL; AB089343; BAC22065.1; -.

GO; GO:0016787; F:hydrolase activity; IEA.

InterPro; IPR002860; GH_BNR.
                                                                                                                                                                                                                                                                                                                      590 FYRSTDGGVTFQPVAAGL------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=M128;
MEDLINE=22359052; PubMed=12374797;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative 124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yaoi K., Mitsuishi Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF02012; BNR; 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=203496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Simi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Geotrichum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8J0D2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
Q8J0D2
                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                       da Silva A.C.K., Ferro J.A., Ferrol C.B., Van Sluys M.A., Almeida N.F.,
Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
Alves L.M. Coutinho L.L., Cursino-Sancos J.R., El-Dorry H.,
Raria J.B., Ferreira A.S., Ferreira R.C.C., Ferro M.I.T.,
A Formighieri B.F., Franco M.C., Greggio C.C., Gruber A., Lemos M.V.F.,
A Formighieri B.F., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Martins B.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
A Martins B.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
A Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira B.C., Tezza R.I.D.,
M. Stindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
M. Schubal J.C., Kitajima J.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SERIOR SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERVICIONE SERV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATSGPYQWRSVAIGGGGFVTGVLFHPAERGLAYARTDVGGAYRWDAQAQQWTALTDWLGA 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PWLIFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLIKWDSGGQIHI 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APORPLOWWFODRGLEETVPLDLLSPMAGAHLLSALGDIDGFRHDELDRA-QLOYAGPRL 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NOLGRANGERLAVDPHDGRVLLLGSRDA-GLWRSDDRGAHWAKVASFPDAALAGATARNH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 TGYQSDIQGVVWVAFDKSSSSLGQASKT1FVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RPSHMAGGSDGH-WYLSYGDQPGPDLMAGGALWKFTPAQGRWREISPIPQ-PASGDGFGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 AP-----MVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           78;
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE012276; AAM41043.1; -.
GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
GO; GO:0006457; P:protein folding; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR001179; FKBP PPlase.
InterPro; IPR002860; GH_BÑR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                751 AA; 81140 MW; 7378B6AA0B5698A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative 121; Mismatches 295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28.4%; Score 1145; DB 35.0%; Pred. No. 2e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROSITE; PS00453; FKBP PPIASE 1; 1.
PROSITE; PS00639; THIOL PROTEASE HIS; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR000169; SHProt_acsite.
Pfam; PF02012; BNR; 7.
                                (anthomonadaceae; Xanthomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 417:459-463 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Complete proteome
SEQUENCE 751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Simi
Matches 266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
```

셤 ò 요 ò g ò Q ò g ò 셤 ò g ò

```
58 GASWIPLWDWAVYPILNKYYSYNAAPAPW----IGPNIVDVTPGNLOIGWMMESLSIDPF 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64 VAGIV---GQPATYYFGTPGGGLWKTINAGRTWQPIFDSVNVASIG-----ALALAPSNP 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 NKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGE----RLAV 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291 TDTANDYFGYSGLTIDROHPNTIMVATOISWWPDTIIFRSTDGGATWTRIWDWTSYPNRS 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       409 --LTKW---DSGGOIHIAPMVKGLEETAVNDLISP-----PSGAPLISALGDLG--- 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SPTHADVTAVPSTIFT-----SPVFTTGTSVDYAEL 483
 1 KYSIDTGAWTDITPVSGSDL---FFGFGGLALDTQKNGTVMVAALNSWWPDGQIFRSTNG 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----IIYAA----TGTVGFGP----GAARPNGPSAFIYRSNDEGSTWQLVGGAG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241 IPHKGVFDPVNHVLYIATSNTG-GPYDGSSGDVWKFSVTSG-----TWTRISPVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          382 ATWTEWFNQPTGEMYHV------VTDTAFPYKLYASQQDSGSVAVTSRSDFGQIT
                               GATWIRIWDWISYPNRSLRYVLDISAEPWLIFGVOPNPPVPSP---KLGWMDEAMAIDPF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VDGIVFNEGAPGILYVRIDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DPNNDNILY-----FGAPSGKGLWRSTDSGATWSQMTNFPDVGTYI----ANPTDTTGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             184 QSDIQCVVWVAFDKSSSSLGQASKTIFVGVADPNNP---VFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    V-LGSGYF--SRVFVNPKNADDVFVM-----QTATYRSQDGGKTFA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              351 LRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATND--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                             Eck J., Schleper C.;
"Acidobacteria form a coherent but highly diverse group within the
bacterial domain and exchange genes with Rhizobiales: Evidences from
                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
Quaiser A., Ochsenreiter T., Lanz C., Schuster S.C., Treusch A.H.,
Eck J., Schleper C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1064;
                                                                                                        Indels
                                                                                       391 NSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLIS
                                                                                                                                                                                                                                                                                                                                                                                                                       population genomics.";
Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
EMBL; AY281358; AAP58622.1; -.
Hypotherian protein.
SEQUENCE 1064 AA; 114845 MW; 36999BF7A7E4ADDA CRC64;
                                                                                                                                                                                                                         (TrEMBLrel. 25, Created)
(TrEMBLrel. 25, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                % Match 9.1%; Score 368.5; DB 2; Local Similarity 23.0%; Pred. No. 3.3e-12; les 171; Conservative 94; Mismatches 243;
                                                                                                                                                                                                                                                                                                        samples
                                                                                                                                                                                                 Ą
                                                                                                                                                                                                                                                                                  uncultured Acidobacteria bacterium.
Bacteria; Acidobacteria; environmental
NCBI_TaxID=171953;
                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                            01-OCT-2003 (TrEMBLrel
01-OCT-2003 (TrEMBLrel
01-OCT-2003 (TrEMBLrel
Hypothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13
                               334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                  RESULT 15
                                                                                                                                                                                   Q7X2U2
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                         엽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                               ŝ
                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                        379
                                                                                                                     417
                                                                                                                                                                  528
                                                                                                                                                                                                                                                                         544
                                                                                                                                                                                                                                                                                                     588
                                                                                                                                                                                                                                                                                                                                   598
                                                                                                                                                                                                                                                                                                                                                                644
                                                                                                                                                                                                                                                                                                                                                                                           TFQPVAAGLPSSGAVGVMFHAVPGKEGDLWL-AASSGLYHSTNGGSSWSAI-TGVSSAVN 656
                                                                                                                                                                                                                                                                                                                                                                                                                          702
                                                                                                                                                                                                                                                                                                                                                                                                                                                        710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  333
                                                                                                                                                    431
                                                                                                                                                                                                            432 VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRA 491
TGAFPDKKPASIAPQPMKVALTP--NFLYVTYADYPGPWGVTFGEVWRQNRTSGAWDDIT 299
                                                                                                                                                                                                                             GSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTT---GGTVAASADGSRFVWAP--G
                                                                                                                                                                                                                                                                                          300 PRVGNSSPAPYNNQTFPAGGFCGLSVDATNPNRLVVIT-LDRDPGPALDSIYLSTDAGAT
                                                                                                                   359 WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTPVPWLDFNNGPQWGGYGAPHGTPGLTKFG
                                                                                                                                                                                                                                                                                                                                 DPGQPVVYAVGFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALSNGTFYRSTDGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WTRIWDWTS-----YPNRSLRYVLDISAEPWLTFGVQP-----NPPVPSP---KLG
                                                                                                                                                  WMDEAMAIDPFNSDRMLYGTGATLYATNDLTK----WDSGGQIHIAP----MVKGLEETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      VGFGKS----APGSSYPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI
                            ----SPVPSTDTANDYFGYSGLTIDROHPNTIMVATQISWWPDTI---IFRSTDGGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    703 VGAPKSKSDGKKASAPSAVFINGTDKPGSDIGLYRSDDNGSTWTRVNDQEHNYSG-PTMI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bisporus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Agaricus bisporus (Common mushroom).
Sakaryota, Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales: Agaricaceae; Agaricus.
NCBL_TaxID=5341;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Molecular analysis of the cellulolytic genes in Agaricus bispor
Thesis (2001), Department of Microbiology, University of London,
London, United Kingdom.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          196 196
196 AA; 21353 MW; D9308B2B09935FA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Putative endoglucanase (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10.1%; Score 409; DB 3;
llarity 45.5%; Pred. No. 2.1e-15;
Conservative 30; Mismatches 51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         196 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EADPKVYGRVYLGTNGRGIVYADL 785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TGDHANLRRVYIGTNGRGIVYGDI 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AJ534359; CAD58875.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Morales-Almora P.;
Submitted (DEC-2002) to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Simi
Matches 76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Morales P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON TER
NON TER
SEQUENCE
                                                                                                                                                  380 1
                                                                                                                                                                                                                                                                                                                                 545
                                                                                                                                                                                                                                                                                                                                                               589
                                                                                                                                                                                                                                                                                                                                                                                            , 669
                                                                                                                                                                                                                                                                      492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              762
                                                                                       337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8J1H7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8J1H7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                        셤
                                                                                                              g
                                                                                                                                                                           엄
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                          ò
                                                                                 8
                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
```

37;

78

183

256

Search completed: May 11, 2004, 12:09:26 Job time : 90.3468 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 12:04:16 ; Search time 36.3141 Seconds (without alignments) 1052.023 Million cell updates/sec

US-09-917-376-3 4036 1 ATTQPYTWSNVAIGGGGFVD......XIGTNGRGIVYGDIGGAPSG 740 Title: Perfect score: Sequence:

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

389414 Total number of hits satisfying chosen parameters:

389414 segs, 51625971 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Issued_Patents AA:*
1: /cgn2_6/ptodata/2/laa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/laa/5A_COMB.pep:*
3: /cgn2_6/ptodata/2/laa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/laa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/laa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/laa/PCTUS_COMB.pep:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		ok			SUMMARIES		
Result No.	Score	Query	Query Match Length	DB	ID	Description	
, , , ,	219.5	5.4	2736	4	US-09-252-991A-30227	Sequence 30227, A	
7	181.5	4.5	3892	4	US-09-328-352-5503	Sequence 5503, Ap	
m	162.5	4.0	772	~	US-08-410-784A-5	'n	
4	158.5	3.9	774	٣	US-09-346-237-8		
5	157.5	3.9	688	4	US-09-489-039A-9813	98	
9	154	3.8	1548	4	US-09-252-991A-22301	223	
7	152.5	3.8	1751	٣	US-09-136-574A-44	44	
80	152	3.8	3472	4	US-09-408-020-4	4	
6	151.5	3.8	776	m	US-09-346-237-4	4,	
10	148.5	3.7	750	9	5457037-3		
11	148.5	3.7	751	9	5457037-5	Patent No. 5457037	
12	148.5	3.7	176	ო	US-09-346-237-7	7,	
13	147	3.6	3623	4	US-09-341-461-2	Sequence 2, Appli	
14	146	3.6		4	US-09-640-419C-28	Sequence 28, Appl	
15	143	3.5		ო	US-08-362-525-22	Sequence 22, Appl	
16	143	3.5		m	US-08-971-692-15	Sequence 15, Appl	
17	142	3.5	_	4	US-09-134-000C-5130	Sequence 5130, Ap	
18	141	3.5		m	US-09-296-284-25	Sequence 25, Appl	
19	141	3.5		m	US-09-296-284-4	Sequence 4, Appli	
20	140.5	3.5		4	US-09-252-991A-21962	Sequence 21962, A	
21	140.5	3.5	_	٣	US-09-352-159-25	Sequence 25, Appl	
22	140.5	3.5	-	m	US-09-352-168-25	Sequence 25, Appl	
23	139	3.4		4	US-09-328-352-7515	Sequence 7515, Ap	
24	137.5	3.4		4	US-09-252-991A-30579	Sequence 30579, A	
25	137	3.4	619	4	US-09-252-991A-26352		
26	136	3.4	2628	~	US-08-570-311-14	Sequence 14, Appl	
27	135.5	3.4	1205	М	US-09-352-159-29		

Sequence 29, Appl	Sequence 6045, Ap	Patent No. 5268290	Sequence 8, Appli	Sequence 8, Appli	Sequence 15, Appl	Sequence 6, Appli	Sequence 6, Appli	Sequence 6, Appli	Sequence 258, App	Sequence 262, App	Sequence 266, App	Sequence 5999, Ap	Sequence 4, Appli	Sequence 3, Appli	Seguence 3, Appli	Sequence 5, Appli	Sequence 5, Appli
US-09-352-168-29	US-09-328-352-6045	5268290-2	US-08-386-727-8	US-08-600-452A-8	US-09-540-245A-15	US-08-542-003-6	US-08-322-760A-6	US-09-236-949-6	US-09-071-035-258	US-09-071-035-262	US-09-071-035-266	US-09-134-000C-5999	US-08-325-267A-4	US-09-397-885-3	US-09-969-362-3	US-09-147-236-5	US-09-522-474-5
m	4	ø	-	~	٣	N	~	4	4	4	4	4	Н	4	4	4	4
1205	1295	626	998	998	1395	1026	1026	1026	1638	1638	1638	1747	862	1277	1277	1326	1326
3.4	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.5	3.5	3.5	3.5	3.5	3.2	3.5	3.2	3.2
135.5	133.5	132.5	132.5	132.5	132.5	131.5	131.5	131.5	131	131	131	131	130.5	130.5	130.5	130.5	130.5

ALIGNMENTS

~

엄

ð

g

8

g

⋩

g

ò

요

ò

ò

g

ò g ò g

à g g

Š

임

ò

Ωp

ò

```
1137 AVVIIDTTAPNAPVLDPINATDP------VSGTAEAGST-VTVTYPDGTT--ATVVAG 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                .239 TNDSTPALIGTVNDPTATVVVN--VDGTDYPAVNNGDG-TWTLADNTLPVLADGPHT-IT 1294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1295 VTATDAAGNAGTDTAVVTIDTTAPNA--PVLDPINATDPVSGTAEAGSTVTVTYPDGTTA 1352
                                                                                                                                                                                                                                        319
                                                                                                                                                                                            855
                                                                                                                                                                                                                                                                                                                                     365
                                                                                                                                                                                                                                                                                                                                                                                 913 NGQTYTATVDSTAG--TWT----VSVPGSDLTADADKTIDAKVTFTDAAGNSSSVNDTH 965
                                                                                                                                                                                                                                                                                                                                                                                                                              366 ----GVQPNPPVPSPKLGWMDEAMAIDPFNSD-----RMLY--GTGATLYATND--- 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --TSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAAS 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                534 ADGSRFVWAPGDPGQPV----VYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGT 589
697 NVTVTGVLKNVPADAANTVVTVVINGQTYTATVDSTAG------TWTVSVPGSDLTAD 748
                                                                                          749 ADKTIDAKVTFTDAAGNSSSVNDTQTYTIDTTAPDAPVINPVNGTDPITGTAEPGSTVTV 808
                                                                                                                                                                                                                                                                                    856 ATVDAVGPNTDGVN--FTVDSVTADNVINASEASGNVTVTGVLKNVPADAANTV-VTVVI 912
                                              -- FVGVADPNNPVFW 222
                                                                                                                                           233 SRDGGATWQAVPGAP-TGF1PHKGVFDPVNHVLY1ATSNTGGPYDGSSGDVWKFSVTSGT 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ITGVSSAVNVG-----FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDC-----GTTW 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             590 FYRSTDG--GVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISPPSGAPLISA-----LGDLGGFTHAD-----VTAVPSTIFTSPVFTTG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --APMVKGLEETAVND---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353 TVVAGTDGSWSVPNPGNLVDGDTVTATATDPAGNTSLPGTGTVSADI 1399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUI, A STARCH DEBRANCHING
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    694 VLI--NDDQHQYGNWGQAITGDHANLRRVYIGTN----GRGIVYGDI
                                                                                                                                                                                                                                                                                                                                     320 SWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF
                                                                                                                                                                                          809 TYPDĆSTTTVVAGPDGTWTVPNPGLNDG-DKVTALATDPAGNP-
                                                                                                                                                                                                                                        282 WTRISPVPSTDTANDYFGYSGLTIDR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----LTKWDSGGQIHI--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 5, Application US/08410784A Patent No. 5912413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-410-784A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     476
                                                                                          g
                                                                                                                               ઠે
                                                                                                                                                                                셤
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 5503, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
FOR THE OF INVENTION: Breton et al.
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
FURRENT APPLICATION NUMBER: US/09/328,352
CURRENT PILLNG DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
                                                                                                                                                                                                                     : | : | | | | : : | 1841 TPVVNPSNGSVIAGTAEAGATVILTDGGGNPIGQVTADGSGN-WSFTPGTPLANGSVINA 1899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1788 NPIGÓA-----TADGSGNWSFTPGTPLANGTVINAVAQDÞAGNTSGPTSTTVDAVAPA 1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1900 LAQDAAGNTSGPASTTVDSVAPATPVLDPSNGTVISGTAEAGATVILTDGG--GNPIRQA 1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2018 GTAEPGVRVILTDGNGNPIGQTLADGSGNWTFTPGTPLANGTVVNAVAQDPAGNTSGPAS 2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54;
                                                                                                                                                                                       ----ISALG-DLGGFTHADV-TAVPSTIFTSPVF--TTGTSV-DYAELNPSIIVRAGSF 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----DRVNPKT-FYALSNGTFYRS-----TDGGVTFQPVAAG 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    607 LPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNG--GSSWSAITGVSSAVNVGFGKSAP 664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VGWNNWGYNGV--VSIAADPINTNKVWAAVGMYTN------SWD-PNDGAILRS 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            598 V-----NGTDPITGTAEPGSTVTV-----TYPNGDTATVVAGPDGSWSVPNPG--LND 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----- VTGAYRSDDCGTT---- 692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         543 TWT-VSVPGSGLVADADKTIDAKVTFTDAAGNSSTVNDTQI---YTLDTA-APAAPVIDP 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               K----GLWRSTDSGATWSQWTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQ 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TIVDIVAPATPVINASN-----GSVITGT-AEVGAKVILIDGNGNPIGETTADGSG 2127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------WVLINDDQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----IGOVTADGSGKWAFTPATPLANGTVINALAQDAAGNNSSPTSATVDSLAPAAPVI
                                                                                                                                                                                                                                                                                  DPSSQPNDRHVAFSTDGGKNW---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----- DNILYFGAPSG
                                                                                          DPFNSDRMLYGT---GATLYATNDLTKWDSGGQIHIAPMVKGLEETAVNDLISPPSGAPL
                                                                                                                                                                                                                                                                                                                                                                               ------VAASADGSRFVWAPGDPGQPV--VYAVGFGNSWAASQGVP-ANAQIRS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TWSNVAIGGGGFVDG-----IVFNEGAPGILYVR-TDIGGMYRWDAANGRWIPLLDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.5%; Score 181.5; DB 4; Length 3892;
21.5%; Pred. No. 0.00047;
iive 90; Mismatches 344; Indels 309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  103 SDOGATWOITPLPFKLGGN--MPGRGMGERLAVDPNN---
RSLRYVLDISAEPWLTFGVQPNPPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Acinetobacter baumannii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    665 GSSYPAVFVVGTIGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 21.5*
Matches 204; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; ORGANISM: AC1NE
US-09-328-352-5503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-328-352-5503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 5503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
```

```
76 INTNKVWAA--------VGMYTNSWDPNDGAILRSSDQGATWQITPL 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      115 PFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYI 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  175 ANPIDITGYQSDIQG------VVWVAFDKSSSSLGQASKTIFVGVADPNNP 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  220 VFWSRDGGATWQAV---PGAPTGFI------PHKGVFDPV---NHVLYIATSNTG 262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          263 -----GPYDGSSGDVWKFSVTSGTWTRISPV------PSTDTANDYFGYSGLTI 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --HPNTIMVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         342 DWTSYPNRSIR-----YVLDISAE-PWLTFGVQPNPPVPS-----PKLGWMDEAM 385
           ---SWAASQG---VPANAQIRSDR---VNPK 580
                                                                                                                                                                                                                                                                                                                                                                                  -----GDSNDSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTD 732
                                                                                                                                                                                                                                  ----ADSNYWNTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        597 TPLMOGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQSNFYTFAQRLIRSARHIPLRPS
                                                                                                                                                          581 TFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTN
                                                                                                                                                                                                                                                                                                          641 GGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVV--GTIGGVTGAYRSDDCGTTWVLIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94 P---TSTLSGQGLG-----GTLYYGYRAWGPNW---PYNASWTKGSSL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.9%; Score 158.5; DB 3; Length 774; 18.8%; Pred. No. 0.0025;
                                                                                                                                                                                                                                  ----SGAV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                      699 DOHOYGNWGQAITGDHANLRRVYIGTNGRGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              733 T----CDWNDGASTFVAPGSETLIG--GAGTTYGQCG 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 18.8%; Pred. No. 0.0025;
Matches 158; Conservative 103; Mismatches 265;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 5629.200-US
CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT APPLICATION NUMBER: PA 1998 00868
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: 60/094,353
EARLIER APPLICATION NUMBER: 60/094,353
EARLIER FILING DATE: 1998-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FRSEEGO for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 8, Application US/09346237A Patent No. 6265197 GENERAL INFORMATION:
                                                                                                                                                                                                            TYPE: PRT
ORGANISM: Favobacterium odoratum
           545 DP----GQPVVYAVGFGN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME/KEY: PEPTIDE
LOCATION: (1)..(774)
OTHER INFORMATION: Isoamylase
                                                                                                                                                                                                                                                                                                                                                                                  684 NYAIAYAINGPSL----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-346-237-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-346-237-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                          Š
                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                          ઠ
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    544 NGANNSQASYGPSDG----GTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEMLSAG 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----FWSRDGGATWQA 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::| | :: | | :: | | :: | | :: | | :: | | :: | | :: | | :: | | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: | :: |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           436 ISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR---- 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --AGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTV---AASADGSRF-VWAPG 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     105 QGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQM 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----DTANDYF-GYSGLTIDROHPNTI---MVATQISWWPDTIIFRSTD 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGATWTRIWDWTSYPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNS 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GFR-----FDLASVLGNSCLNAVHASA-----PNCPNGGYNFDAADSNVAI-----436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               437 NRILREFTVRPAAGGTVWICLRNLGPSAATR--TSWVDSRRVVRVEWSVPRQLRQAQNEL 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51 WIPLLDWVGWNNWGYN---GVVS---IAADPINTNKVWAAVGMYTNSWDPNDGAILRSSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WGP--NWPYASNWGKGSQAGCVSDVDANGDRFNPNKLLLDPYAQEBSQDP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----LNPSNQN----GNVFASAHYRTTDSGI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     165 TNFPDVGTYIANPTDTTG------YQSDIQGVVWVAFDKSSSSLG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    233 VPGAPTG-FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAGGPTAEFQAMVQAFHNAGIKVYMDV-------VYNHTAEGGTWTSSDPTTAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               393 DRMLY-----GTGATLY-----ATNDLTKW-DSGGQIHIAPMVKGLEETAVNDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 293;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 772;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90; Mismatches 272;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---QASKTIFVGVA------DPNNPV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.0%; Score 162.5; DB 19.8%; Pred. No. 0.0012
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: 34,346
REFERENCE/DOCKET NUMBER: 1SU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEXX:
                                                                                                                                                                                                                                                                                                                                                NAME: Heine, Ph.D., Holliday REGISTRATION NUMBER: 34,346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 772 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                    FILING DATE: ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 19.8%
Matches 162; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ANTI-SENSE: NO FRAGMENT TYPE: NO CRIGINAL SOURCE: US-08-410-784A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                292
```

à

ò

ò

8 셤 Š ò

ઠ

Db 382 DPGNPALTGVPPEGQTYTRGTPNVWSAMSYDAKLNLIYLPTGNATPD 428 Qy 297 YFGYSGLTIDRQHPNTINVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLR 352	Db 478 YDLPDGKGGTTPVLVQTSKQGMIFMLNRETGEPVAKVERRPVPAGNVKGERYSPTQ 533 Qy 396 LYGTGATLYATNDLTKWDSGQCHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGG 453 Db 534 PYSVGMPMIGNQTLTESDWMGATPIDLLLCRIQFKEMRHQGVFTPPGEDRSLQFPGSLGG 593	Qy 454 FTHADVTAVPS 464 Db 594 MINGSVSLDPN 604	SEQUENCE 22301 Sequence 22301, Application US/09252991A Patent No. 6551795 Sequence 22301, Application US/09252991A Patent No. 6551795 ABRERAL INFORMATION: REPLICANT: Marc J. Rubenfield et al. TITLE OF INVENTION: AERCGINOSA FOR DIAGNOSTICS AND THERAPEUTICS CURRENT APPLICATION NUMBER: US/09/252,991A CURRENT FILING DATE: 1999-02-18 PRIOR FILING DATE: 1998-02-18 PRIOR FILING DATE: 1998-07-18 PRIOR FILING DATE: 1998-07-27 NUMBER OF SEQ ID NOS: 33142 SEQ ID NO 22301 LENGTH: 1548 TYPE: PRT ORGANISM: Pseudomonas aeruginosa -09-252-991A-22301 Query Match 3.8*; Score 154; DB 4; Length 1548; Best Local Similarity 20.7*; Pred: No. 0.014; Matches 188; Conservative 87; Mismatches 270; Indels 364; Gaps 73 ADPINTNKVWAAVGMYTNSWD	10ZSEDQGATWQTTPLPFKKLGGNMPGGKGMGEKLAVDPNNDN LYFGAFSGKGLMK	DB 3/9 EQSSGGLVFTDQATDWGITFGTPVDESNEM1CGLGPYC 416 QY 154 STDSGATWSQMTNPPDVGTYIANPTDTTGYQSDIQGVVWVAPDKSSSSLGQASKTIFV 211	Db 417 AEPPRWTWAPGNYLGMPAGTAIGEGDLWWCSVDP\$LCIENLGKT 460 Qy 212 GVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDP 249	Db 461 -VVTPQNQLFSVLRTGTGDLDLASAGNLTQWSPYGVYTAGTQAADVA-TGFNQPRGLFN-517 Qy 250 VNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVP 289		QY 290 STDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIF-RSTDGGATWTRIMDWT 344 : : : : : : : Db 575 PSAAVGNWLWRQG-SADREGVPTAWWVNFGSYVRGAEGDAPYLVGFTGFGT 624	Qy 345 364
	0b 446 GGŚGVDLİAEPWĀIĠGNSYQVĠĠPPSGWAEWNGAYRĎVVRQAQNKLGŚVAITTĠ 499 2y 482 ELNPSIIVR-AGSPDPSSQPNDRHVAFSTDGGKNWPQGSEPGGVTTGGTVAASADGS 537		ST2 IRS		Query Maccn 3.9%; Score 15/.5; DH 4; Length 688; Best Local Similarity 22.6%; Pred. No. 0.0025; Matches 111; Conservative 53; Mismatches 184; Indels 143; Gaps 24;	2y 60 WNNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFK 117 : :	225IGDTLYVCTPYSKVLALDVDSGKEKMRYDSKSSPNWQRCRGLGYYADSQA 275IGDTLYVCTPYSKVLALDVDSGKEKMRYDSKSSSPNWQRCRGLGYYADSQA 275 · · · · · · · · · · · · · · · · · · ·	165 TNFPDVGTYIANPTDTTGYQSDIQGVVWVAFD	Db 276 QTAPASGTQPAACSRRLFLPTIDARLIAIDADTGKLCENFGDGGIVDLSVGMGEVKAGYY 335 197 KSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVF 247	b 336 QQTSTPLVAGNVVVVGGRVADNYSTGEPPGVVRAFDVHTGKLAWAW 381 248 DPVNHVLXIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAND 296

8 4 8 6 8 6 8 6 8 6 8 6 8 6 8 6

REGIGTRATION NUMBER: 31,215 REGIGTRATION NUMBER: 31,215 REFERENCE/DOCKET NUMBER: 1997US001/CIP TELECOMMUNICATION INFORMATION: TELEFAX: 215-540-2800 TELEFAX: 215-540-2818 TELEFAX: 215-540-2818 TELEFAX: 215-540-3818 TELEFAX: 215-540-3818 TELEFAX: 215-540-3818 TELEFAX: 215-540-3818 TELEFAX: 215-540-3818 SEQUENCE CHARACTERISTICS: TYPE: amino acids TYPE: amino acids TYPE: amino acids TYPE: amino acids TYPE: amino acids TYPE: anino acids TYPE: anino acids TYPE: anino acids TYPE: anino acids TYPE: anino acids TYPE: anino acids TYPE: anino acids TYPE: protein SEQUENCE DESCRIPTION: SEQ ID NO: 44: US-09-136-574A-44	Query Match 3.8%; Score 152.5; DB 3; Length 1751; Best Local Similarity 18.7%; Pred. No. 0.021; Matches 184; Conservative 101; Mismatches 349; Indels 351; Gaps 49; Qy 6 YTWSNVAIGGGGFVDGIVFNEGAPGILLVVRTDIGGMYRWD 45	46 AANGRWIPLLDWYGWNNWGYNGVYSIAADPINTNKVWAAVGM-YTNSWDPNDGAILRSSD	105 QGATWQITPLPFKLGCMMPGRGMGERLAVDPNNDNI : :	Qy 141 LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANFTDTTGYQSDIQGVVW 192	Qy	482 EIEVESKEGNSQGPNYTEVISYIYNRTGWPPRVTDKLSFKYFIDLTELIQAGY	SGLFIDKQHEVIINVAIQISWWFDIIIFKSIDGGATWIKIWWNISIFAKSIDISM	OY 360 EPWLTFGVQNNPPVPSPKLGWMDEAMALDPFNSRMLYGTGATLYATNDLTKWDSG 415	Qy	Qy 444PGFPHADVTAVP 463 1 1 1 1 1 1 1 1 1 1	Qy 464 STIFTSPVFTTGTS-VDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKN514 :	Qy 515 WFQ GSEPGGVTTGGTVAASADGSRFVWAPGDP 546	QY 547GQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKIFYALSN-GT 589 DD 860 VSATPTPAFTASPVG-GSYWTPSESYGALKVWYANGNLSSPTNVLNPKIKIENVGTTA 916
365FGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGAT 402	521 PGGYTTGGTVAASADGSRFVWAPGDPGOPVYYAVGFGNSWAASQGVPANAQI 572	MLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSD :::	DCGTTWVLINDD : GAGPVWMRAGRD	732 GDIGGAPSG 740 	Jur 7 39-136-574A-44 ayuence 44, Application US/09136574A tent No. 6294366	GENERAL INFORMATION: APPLICANT: Farrington, Graham K. Anderson, Paige	Gibbs, Moreland Bergquist, Peter Daniels, Roy Morgan, Hugh W	Williams, Diane P. TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing Fabrics Using Truncated Cellulase Envine Compositions	NUMBER OF SEQUENCES: 49 CORRESPONDENCE ADDRESS: ADDRESSEE: Howeon and Howson carpen: Carrier House Corrected Carter D O Box 467		COMPUTER READABLE FORM: MEDIUM TYPE: Diskette COMPUTER: IBM Compatible	CURRENT APPLICATION DATA: CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/136,574A FILING DATE: 19-Aug-1998	PRIOR APPLICATION: <unknown> PRIOR APPLICATION DATA: APPLICATION WIMBER: US 08/932,571 FILING DATE: September 19, 1997 ATTORNEY/AGENT INFORMATION:</unknown>

ဖ

g

g ò g ò

ò

```
1380 ----VPSPSGIEFSDDGMRM-----FVTGIGTPGINIFTLSAPFDITLPKHSGSTNIG 1428
                                                                                                                                                                                              1482 IPNSIYTSPDGLSQFVAXDDRIDLYVLGSPNDISSTTEIIPYSLPRP---DPPTGMDFTP 1538
                                                                                                                                                                                                                                                                                  240
                                              422 PMVKGLEETAVNDLISPPSGAPL-----ISALGD----LGGFTH----ADVT-A 461
                                                                                                                                                                                                                                                                                                                                                   548 -- QPVVYAVGFGNSW--AASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPV 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                      604 AAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSA 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87 MYTNSWDPNDGAILRSSDQCATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 183 KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               147 SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               243 IG-VTAVEFLPVOETQNDANDVVPNSDANONYWGYMTENYFSPDRRYAYNKAAGGPTAEF 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG 86
                                                                                                                                                  ---VFTTGTSVDYA---ELNPSIIVRAGSFDPSS----QP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LGQASKTIFVGVAD-----PN---PN----SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                     NDRHVAFSTDGGKNWFQGSEPGGVTTGG-----TVAASADGSRFV-----WAPGDPG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels 291; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 776;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1680 PFDTTLPVHVELHDİĞĞQPAVDLAFAED--ĞRTLLLLAAD 1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         664 P-GSSYPAVEVVGTIGG----VTGAYRSDDCGTTWVLINDD 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 3.8%; Score 151.5; DB 3; Best Local Similarity 20.7%; Pred. No. 0.008; Matches 171; Conservative 89; Mismatches 274;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsan, Allan
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 5629.200-US
CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: PA 1998 00868
EARLIER FILING DATE: 1998-07-02
EARLIER FILING DATE: 1998-07-02
EARLIER FILING DATE: 1998-07-28
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 4
SEG ID NO 4
LENGTH: 776
                                                                                                                                                                                                                                                                                                                                                                                                    QRHELIYPYGASTSLLETVRDGV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Pseudomonas amyloderamosa
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 4, Application US/09346237A
; Patent No. 6265197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; LOCATION: (1)..(776)
; OTHER INFORMATION: Isoamylase
US-09-346-237-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: PEPTIDE
                                                                                                                                                  462
                                                                                                                                                                                                                                                     501
                                                                                        g
                                                                                                                                       ò
                                                                                                                                                                                      g
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                               ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                          |: | | : : | | : : | | : : | | 377 VLAAGQSTKEIRLSIQKGSGSYNQSNDSVRSANSYIENEKVTGYIDDVLVWGREPGRNA 1036
                                                                                                                                                                                                                           : | : | : | : | : | 1096
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DGTVVFLSD-GPTGRVYPYSLNIPFDISSAAPG-----GFVIVPVGVSDIAFSA-DGRNM 1038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1268 MEQVTLDSREGTLRVRLDRTVDVDTVRPYKMWVEDSDGSQTTLANSTLLNAENSNILLFR 1327
614
                                                                                                -----ITGVSSAVNVGFGKSAPGS 666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DGAILRSSDQGATWQITP----LPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGL 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          152 WRSTDSGATWSQMTNFP-DVG-----TYIANP--TDTTGYQSDIQGVVWVAF 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DKSSS------SLGQASKTIFVGVADPNNPVFWS------RDGGATWQAVPG- 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---LTIDRQH------PNTIMVAT---QISWWPDT--------IIFR 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    330 STDGGA------TWTRIW---DWTSYPNR-----SLRYVLDISAEPWLTFGVQ 368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          369 PNPPVPSPK-LGWMDEAMAIDPFNSDRMLYGTGA-----TLYATNDLTKWDSGGOIHIA 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Scaleper, Christa
APPLICANT: Feldman, Robert A.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCORP.002A
CURRENT PELLING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/102,294
PRIOR PILING DATE: 1998-09-29
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 4.
                            -APTGFIPHKGVFDPV----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----YDGSSGD--VWKFSVTSGTWTRISPVPSTDTAN----DYFGYSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 3.8%; Score 152; DB 4; Length 3472; Best Local Similarity 21.0%; Pred. No. 0.058; Matches 172; Conservative 100; Mismatches 244; Indels 304;
                                                                                           VMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA-
                                                                                                                                                                                              667 SYPAVFVVGTIGGVTGA----YRSDDCGTT
                                                                                                                                                                                                                                                                                                                                              1097 PAYIDVKFVKLGANAGGADYYVEIG 1121
                                                                                                                                                                                                                                                                                               712 GDHANLRRVYIGTNGRGI-VYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4, Application US/09408020 Patent No. 6632937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Cenarchaeum symbiosum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 3472
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-408-020-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
```

g

ò

ò

셤

ò

g

à

셤

ò

QQ ò 셤

ઠે

DLASVLGNSCLNGAYTASAPNCPNGGYNFDAADSNVALNRILREFTVRPAA 451 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEETAVNDLISPPSGAPLI 445	\$ 8 &	202 LGGASKIIRVOKADI
: NGLFRDSLRQAQNELG	a & a	276 QAMVQAFHNAGIKVYMDVVYNHTAEGGTWISSDPTATIYSWRGLDN 322 292DTANDYF-GYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIW 341 323 TTYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTMGVDGFRF 374
SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGS 493 :	ò qa	342 DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRML 396
FDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRF-VWAPCD 545 :	& 43	397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKQLEETAVNDLISPPSGAPLI 445
PGQPVVYAVGFGNSWAASQGVPANAQIRSDRV 577	ك 8	446 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGS 493
NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGD 626 :: :	ර් අ	494 FDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTCGTVAASADGSRF-VWAPGD 545
LWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPG 665	ò a	546 PGQPVVYAVGFGNV-SWAASGGVPANAQIRSDRV 577
SSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710	ර් යි	578 NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGD 626
	λ δ	627 LWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFGKSAPG 665
FARENT NO. 352/03/ APPLICANT: TOGNONI, ANGELO; CARRERA, PAOLO; CAMERINI, BARBARA; GALLI, GIULIANO; LUCCHESE, GIUSEPPE; GRANDI, GUIDO; DI GENNARO, CARLO TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE ; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME MIMBED OF CEDITEMPE.	δ qa	666 SSYPAVFVVGIIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710
CURRENT APPLICATION NUMBER: US/08/1,797 APPLICATION NUMBER: US/08/1,797 FILING APPLICATION DATA: APPLICATION NUMBER: 749,621 FILING DATE: 09-JAN-1993 APPLICATION NUMBER: 224,114 FILING DATE: 19-AUG-1998 ID NO:3: LENGTH: 750 37-3 TY*; Score 148.5; DB 6; Length 750; t. Local Similarity 20.7*; Pred. NO. 0.013; t. Local Similarity 20.7*; Pred. NO. 0.013; ches 171; Conservative 88; Mismatches 275; Indels 291; Gaps 47; Ches 171; Conservative Resemple Length Resembly Conservative Resemple Length Resembly Conservative Resemple Length Resemple Length Resembly Conservative Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resemple Length Resempl	RESULT 11 5457037-5 ; Patent APPLIO ; GALLI, GII ; TITLE ; RIZHNE AI ; RUZHNE AI ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRIOR ; PRI	PACHLE 11 457037-5 Patent No. 5457037 Patent No. 5457037 Patent No. 5457037 Patent No. 5457037 BAPLICANT: TOGNONI, ANGELO; CARRERA, PAOLO; CAMERINI, BARBARA; GALLI, GIULIANO; LUCCHESE, GIUSEPPE, GRANDI, GUIDO; DI GENNARO, CARLO TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME NUMBER OF SEQUENCES: 7 CURRENT APPLICATION DATA: APPLICATION NUMBER: 18961,797 FILING DATE: 19-ANG-1991 APPLICATION NUMBER: 749,621 FILING DATE: 19-AUG-1991 APPLICATION NUMBER: 25-JUL-1988 SEQ ID NO:5: LENGTH: 751
87 MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP 146	Query M Best Lo Matches	y Match 3.7%; Score 148.5; DB 6; Length 751; Local Similarity 20.7%; Pred. No. 0.013; hes 171; Conservative 88; Mismatches 275; Indels 291; Gaps 47;

α

```
g
                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                            ठे
                                                                                                                                                                                                                                                                                             ò
                                                  146
                                                                         157
                                                                                                    201
                                                                                                                            217
                                                                                                                                                     240
                                                                                                                                                                              276
                                                                                                                                                                                                       291
                                                                                                                                                                                                                             277 QAMVQAFHNAGIKVYMDV------VYNHTAEGGTWTSSDPTTATIYSWRGLDN 323
                                                                                                                                                                                                                                                      -----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                        DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML 396
                                                                                                                                                                                                                                                                                                                                                            445
                                                                                                                                                                                                                                                                                                                                                                                    477
                                                                                                                                                                                                                                                                                                                                                                                                             493
                                                                                                                                                                                                                                                                                                                                                                                                                                     531
                                                                                                                                                                                                                                                                                                                                                                                                                                                               545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQ---SNFYTFAQRLIAFRKAHPAL 633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           726
                       71 GITGAVY-----YGYRAWGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY 122
                                                                                                                                                                                                                                                                                                                               ------PNCPNGGYNFDAADSNVAINRILREFTVRPAA 426
 86
                                                                                                                    LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                       : | | | : | : | : | : | TTYYELTSGNQYFYDNTGIGANFNTYNTVAQNLIVDSLAYWANTM---GVDGFR-----F
                                                                                                                                                                                                                                                                                                                                                                                                                             YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP
GAPGILYVRIDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG
                                                                           -----NGNVFASGASYR----TTDSGIY--AP
                                                                                                 SCKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                     LGQASKTIFVGVAD-----PN---PN----SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                                                                                                                            --TAVNDLISPPSGAPLI
                                                                                                                                                                                                                                                                                                                                                                                  427 GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL----GSMTI
                                                                                                                                                                                                                                                                                                                                                                                                          446 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVR-----AGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YGPS-----DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEMLSAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P----GQPVVYAVGFGN------SWAASQGVPANAQIRSDR-----V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --APG
                                               87 MYTNSWDPNDGAILRSSDQGATWQITPLPPKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                    241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                              FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRF-VWAPGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------GSSWSAIT----GVSSAVNVGFGKSAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --LGDSNSIXVAYNGWSSSVTFTLPAPPSGTQWYRVTDTCDWNDGASTFV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SE-----YGQCGQSL 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVG-----
                                                                                                                                                                                                                                                                                                                                                         YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE
                                                                         AQEVSQDP----LNPSNQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 7, Application US/09346237A Patent No. 6265197 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                DLASVLGNSCLNGAYTASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   627 LWLAASSGLYHSTNG-
                                                                        123
                                                                                                                            158
                                                                                                                                                                                                                                                                                                                                                         397
                                                                                                                                                                                                                                                                                                                                                                                                                                   478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      577
                                                                                                 147
                                                                                                                                                                            218
                                                                                                                                                                                                                                                                                                        342
                                                                                                                                                                                                                                                                                                                                376
                                                                                                                                                                                                                                                                                                                                                                                                                                                              494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 578
                                                                                                                                                   202
                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9/9
                                                                                                                                                                                                                                                                               324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    999
                                                                                                                                               8
                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                     g
                                                                   a
                                                                                               ò
                                                                                                                         유
                                                                                                                                                                                                  ð
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
ò
```

```
ò
                                                                   d
                                                                                                                 à
                                                                                                                                                            요
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsen, Allan
FILLE OF INVENTION: Starch Debranching Enzymes
FILLE REFERENCE: 5529.200-US
CURRENT APPLICATION NUMBER: US/09/346,237A
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: PA 1998 00868
EARLIER APPLICATION NUMBER: 60/094,353
EARLIER PILING DATE: 1998-07-28
```

```
47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              503 YVTQDANDFSGS-----SNLFQSSGRSPWNSINFIDVHDGMTLKDVYSCNGANNSQAWP 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              659 RPSSWYSGSOLTWY------OP-----SGAVADSNYWNNTSNYALAYAINGPS-- 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----NGNVFASGASYR----TTDSGIY--AP 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QAMVQAFHNAGIKVYMDV-------VYNHTAEGGTWTSSDFTTATIYSWRGLDN 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----DTANDYF-GYSGLTIDRQHPNTI---MVATQISWWPDTIIFRSTDGGATWTRIW 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DWTS-YPNRSLRYVLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAID----PFNSDRML, 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYABLNPSIIVR------AGS 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            602 PLMQGGDEYLRTLQCNNNAYNLDSSANWLTYSWTTDQ---SNFYTFAQRLIAFRKAHPAL 658
                                                                                                                                                                                                                                                                                                                                                                                                                                          96 GITGAVY-----YGYRAMGPNWPYASNWGKGSQAGFVSDVDANGDRFNPNKLLLDPY 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DLASVLGNSCLNGAYTASA-----PNCPNGGYNFDAADSNVAINRILREFTVRPAA 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----APG 751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGKGLWRSTDSGATWSQMTNPPDV--GTYIANPTD-TTGYQSDIQGVVWVAFDKSS--SS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            183 KGVVLVPSTQSTGTKPTRAQKDDVIYEVHVRGFTEQDTSIPAQYRGTYYGAGLKASYLAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LGQASKTIFVGVAD------PN---NPVFW------SRDGGATWQAVPGAPTG-F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           349 TIYXELISGNQYFYDNIGIGANFNIYNIVAQNLIVDSLAYWANIM---GVDGFR----F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           557 YGPS------DGGTSTNYSWDQGMSAG---TGAAVDQRRAARTGMAFEMLSAGT
                                                                                                                                                                                                                                                                                                                                                                                                 GAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGYNGVVSIAADPINTNKVWAAVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LG-VTAVEFLPVQETQNDANDVVPNSDANQNYWGYMTENYFSPDRRYAYNKAAGGPTAEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              397 YGTGATLYATNDLTKWDSGGQIH-IAPMVKGLEE-----TAVNDLISPPSGAPLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FDPSSQPNDRHVAFSTDGGK----NWFQGSEPGGVTTGGTV---AASADGSRF-VWAPGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P----GQPVVYAVGFGN-----SWAASQGVPANAQIRSDR-----V
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----VMFHAVPGKEGD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GSSWSAIT-----GVSSAVNVGFGKSAPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GSMTI
                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPST----
                                                                                                                                                                                                                                                                                                                                                291;
                                                                                                                                                                                                                                                                                                Length 776;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGSGLDLFA----EPWAIGGNSYQLGGFPQGWSEWNGLFRDSLRQAQNEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SE-----YGQCGQSL 770
                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NPKTFYALSNGTFYRSTDGGVTFQPVAAGLPSSGAVG---
                                                                                                                                                                                                                                                                                             3.7%; Score 148.5; DB 3;
20.7%; Pred. No. 0.013;
ive 88; Mismatches 275;
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSEQ for Windows Version 3.0
EEQ ID NO 7
LENGTH: 776
                                                                                              TYPE: PRT
ORGANISM: Pseudomonas species SMP1
FEATURE:
NAME/KEY: PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             148 AQEVSODP----LNPSNQ-
                                                                                                                                                                                          CONTION: (1)...(776)
COTHER INFORMATION: Isoamylase
US-09-346-237-7
                                                                                                                                                                                                                                                                                                                       Best Local Similarity 20.7
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                627 LWLAASSGLYHSTNG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                494
                                                                                                                                                                                                                                                                                                                                                                                                 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 546
                                                                                                                                                                                                                                                                                                Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             401
```

```
2337
2243 QGAECIWILEAPPGRSIQLQFEDQFNIEDTPNCSVSYLELRDGANSNARLVSKLCGGT-- 2300
                                                                                                                                                                                                                                         969
                                                                                                                                                                                                             697 NDD------QHQYGWWGQAITGDHANLRRVYIGTNG-------RGIVYGDIG 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGNANLFLMNPAGMIFGPNASINVPGDFVVTTGSAIGFG------NDQWFQVFSDNDYN 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 279 GLAVQPNGSV----TTNGTNALVSPLPGSVTISGNVDASGKSTNISSGGQVAIAGDQIAV 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       335 OGATVDVSGNGGGGT------VRIGGDFQGQ--LT-----LPNASQTLI 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  431 -AVNDLISPPSGAPLISALGDLGGF-----THADVTAVPSTIFTSPVFTTGTSVD 479
                                                                             2301 -----LPGS--------WVSSRERIYLKFGTDGGSSYMGFKAKYSIASCG
                                         601 QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLY--HSTNGGSSWSAITGVSSAVNVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  119 GGN-----MPGRGM--GERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        172 TYIANPIDTTGYQSDIQGVVWVAFDKSSSL----GQASKT-----IFVGVADPNNPVFW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                172 ALIGNPSQ------FAFDLANPGLIINAGDLSVTEGKNLTFLAGNIVNTGSLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SRDGGATWQAVPGAPTGFI PHKG-----VFDPVNH-----VLYIATSNTGGPYD---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDROHPNTIMVA-TQISW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             322 WPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLIFGVQPNPPVPSPKLGWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           371 DSNSVV---KADALLTGNGGTVIVWADDSTRFSGNISAQGGTMGGN------GGFVETS
                                                                                                                             659 FGKSAPGS----SYPA----VFVVGTIGGVTGAY-----RSDDCGTTWVLI
                                                                                                                                                                   2338 GTVSGDSGVIESIGYPTLPYANNVFCQWPIRGLPGGYLTLSFEDFNLQSSPGCTKDFVEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Bidney, Dennis L
APPLICANT: H, Xu
APPLICANT: H, Xu
APPLICANT: Lu, Guihua
ITILE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
ITILE OF INVENTION: DEFENSE 1990/640,419C
CURRENT PILING DATE: 2000-08-17
PRIOR PILING DATE: 1999-08-18
PRIOR APPLICATION NUMBER: 60/149,656
PRIOR FILING DATE: 1999-08-18
PRIOR FILING DATE: 2000-05-23
NUMBER OF SEQ ID NOS: 28
SEQ ID NOS: 28
SEQ ID NO 28
LENGTH: 1749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEAMAIDPFNSDRMLYGTGATLYA-TNDLTKWDS------GGQIHIAPMVKGLEET-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      USE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3.6%; Score 146; DB 4; Length 1749; 21.8%; Pred. No. 0.062; tive 81; Mismatches 306; Indels 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 28, Application US/09640419C Patent No. 6630615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Synechocystis PCC6803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3.6
Best Local Similarity 21.8
Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                         | |:|
G-PTG 2461
                                                                                                                                                                                                                                                                                                   736 GAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
US-09-640-419C-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-640-419C-28
                                                                                                                                                                                                                                                                                                                                          2458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    219
                                         8 6
                                                                                                              ઠ
                                                                                                                                                          g
                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --GNEMFVQFISDSSNGGQGFKIRYEAKSLACGGTVYIGDADSDGYLTSPNYPAN---YP 2242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----LSLD------IEPQQSCNYDKLIVKDGDS-------D 2045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1768 ----IASSPGNRLQLSFLSFNLENSLNCNKDFVEIREGNATGGLIGRYCGNSLPGN-- 1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1982 MVTGDTPVGIFSP------GWPREYANG-----ADCIWIIYAP-DSTVELNI 2021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPS 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 I----PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAND 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    297 YFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLD 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----ATLYATNDLTKWDSGGQIHIAPMVKGLEE------TAVNDLI----S 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               498 SQPNDRHVAF---STDGGKNWFQGSEPGGVTTGGTV-AASADGSRFVWAPGDPGQPVVYA 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   554 VGFGNSWAASQGVPANAQIR-----SDRVNPKTFY----ALSNGTFYRSTDGGVTF 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99 ILRSSDQGATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRST--- 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DSGATW-----VGTYIANPTDTT 181
             Sequence 2, Application US/09341461
; Sequence 2, Application US/09341461
; Patent No. 6586389
; GENERAL INFORMATION:
   APPLICANT: Hammond, Timothy G.
   APPLICANT: Hammond, Timothy G.
   APPLICANT: Verroust, Pierre J.
   APPLICANT: Verroust, Pierre J.
   TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
   TITLE OF INVENTION: and Uses Thereof
   FILE REFERENCE: D6148
   CURRENT FILING DATE: 2000-07-20
   PRIOR APPLICATION NUMBER: PCT/US99/01259
   PRIOR FILING DATE: 1999-01-21
   NUMBER OF SEQ ID NOS: 40
   SEQ ID NO 2
   LENGTH: 3623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 YTWSNVAIGGGGF-VDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----INTNK-------VWAAVGMYTNSWDPNDGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182 GYQSDIQGVVW-VAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2100 RGVITSPKYPDTYLPNLNCSWGVLVQTGLTIAV-GFEQPFQIQNRDSFCSQGDYLVLRNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 3.6%; Score 147; DB 4; Length 3623;
Best Local Similarity 20.0%; Pred. No. 0.14;
Matches 181; Conservative 79; Mismatches 301; Indels 344; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                        ; FEATURE:
; OTHER INFORMATION: amino acid sequence of rat cubilin protein
US-09-341-461-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1159 PDNGSPPLGPSGRNGRFCG---MYAPSTLFTS-----
                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        401
                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

```
654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TSNTGGPYD-----GSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTID-- 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --DISAEPWL-TFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKW 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----IVVKTPTTAISSSLSSSSGQITSSITSSRPIITPFYPSNG-TSVISSSV 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   634 -GLYHSTNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVF--VVGTIGGVTGAYRSDDCG 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 DSYGNWGCKGMGACS----NSQGIAYWSTDLFGFYTTPTNVTLEMTGYFLPPQTGSYTFK 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PNDGAILRSSDQGATW-----QITPLPFKLGGNMPGRGMGERLAVDPNNDNIL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 YFGA----PSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDK 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIA 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---TSTSTEMSTVTGTNGLPTDETV-- 444
                                                                                                                                                                                                                                                                                                                                                                                                                   66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGTNGVPTDETVIVIRTPTSEGLI - - - STTTEPWTGTFTSTSTEVTT - ITGTNGQPTDET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           354 VIVIRTPTSEGLISTTTEPWTGTFTSTSTE----MTTVTGTNGQPTDETVIVIRTPTSEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----VTSCESHVCTESISPAIVSTATV-----TVSGVTTEYTTWCPISTTETTKQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       655 KGTTEQTTETTKQTTVVTISSCESDVCSKTAS----PAIVSTSTATINGVTTEY----
                                                                                                                                                                                                                                                                                                                                                                                                                53 TYSNAAYMAYGYA----SKTKLGSVGGQTDISIDYN-----IPCVSSSGTFPCPQE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    156 FATVDDSAİL -- SVGGATAFNCCAQQQPPITSTNFTIDGIKPWGG-----SLPPNIEGTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        209 YMYAGYYYPM-KVVYSNAVSWGTLPISVTLPD-GTTVSD------DFEGYVY-SFD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --DDLSQSNCT----VPDPSN-----YAVSTTTTTTEPWTGTFTSTEMTTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DSGGQIHIAPMVKGLEETAVNDLISPPSGA------PLISALGDLGGFTHADVTAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       463 PSTIFTSPVFTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       494 ISSSVTSSLFTSSPVISSSVISSSTTTSTSIFSESSKSSVIPTSSSTSGSSE-SETSSAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53 GVTTGGTVAASADGSRFVWAPGDPGOPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  553 SVSSSSFISSESSKSP---TYSSSSLPLVTSATTSQETASSL-PPATTTKTSEQTTLVT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          583 YALSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPG--KEGDLWLAASS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----RQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL----
                                                                                                                                                                                                                                                  3.5%; Score 143; DB 3; Length 894;
Similarity 20.5%; Pred. No. 0.041;
51; Conservative 94; Mismatches 308; Indels 224;
                                                                                                                                                                                                                                                                                                                                                          7 TWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG-
                                                                                                                                                                                                                                               Length 894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 --WNNWGYNGVVSIAADPINTNKV--WAA--VGMYTNSWD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: May 11, 2004, 12:11:20 Job time: 40.3141 secs
                                 25:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         410 LVTTTTEPWTGTF-----
                                                                                    894 amino acids
amino acid
6714627 CUSH
NN FOR SEQ ID NO:
CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                           Matches 161; Conservative
                                                                                                                                                             , MOLECULE TYPE: protein US-08-362-525-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TIWALIN 697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TTWCPIS 711
                                                                                                                                             linear
                                                                                                                                          TOPOLOGY:
                                                                                    LENGTH:
                            INFORMATION
SEQUENCE C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                608
        TELEX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              691
                                                                                                                                                                                                                                                  Query Match
Best Local $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 705
```

```
May 11, 2004, 12:09:36 ; Search time 86.3468 Seconds (without alignments) 2378.773 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                             1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgn2_6/ptodata/2/pubpaa/USO7_PUBCOMB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO7_PUBCOMB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO7_NEW_PUB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO7_NEW_PUB.pep:*
/cgn2_6/ptodata/2/pubpaa/PSO7_SUBCOMB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
/cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        todata/2/pubpaa/US09A PUBCOMB.pep:*
ptodata/2/pubpaa/US09E PUBCOMB.pep:*
ptodata/2/pubpaa/US09E PUBCOMB.pep:
ptodata/2/pubpaa/US09E PUBCOMB.pep:*
ptodata/2/pubpaa/US10A PUBCOMB.pep:*
ptodata/2/pubpaa/US10A PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /cgni__6/prodata//2/pubpaa/US10C_PUBCOMB.ppp./cgni__6/ptodata/2/pubpaa/US10 NEW PUB.ppp:*/cgni__6/ptodata/2/pubpaa/US60_NEW PUB.ppp:*/cgni__6/ptodata/2/pubpaa/US60_NEW PUB.ppp:*/cgni__6/ptodata/2/pubpaa/US60_NEW PUB.ppp:*
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1140673 segs, 277566755 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Published Applications AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                 OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                     US-09-917-376-3
4036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cgn2_6/
cgn2_6/
                                                                                                                                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB
Maximum DB
                                                                                                                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                                                                                                         Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Ouery Natch Length DB ID Description No. Score Match Length DB ID Description 1 4036 100.0 740 12 US-09-917-376-3 Sequence 3, Appli 2 4036 100.0 740 14 US-10-155-400-3 Sequence 6, Appli 2 4036 100.0 740 14 US-10-155-400-3 Sequence 7, Appli 2 4036 100.0 740 14 US-10-155-400-6 Sequence 1, Appli 2 4036 100.0 957 14 US-10-155-400-6 Sequence 1, Appli 2 4036 100.0 957 14 US-10-155-400-1 Sequence 1, Appli 3 100.0 957 14 US-10-155-400-1 Sequence 1, Appli 3 100.0 957 14 US-10-155-400-1 Sequence 1, Appli 1 1573.5 39.0 818 14 US-10-155-400-7 Sequence 7, Appli 1 1 1573.5 39.0 818 14 US-10-156-761-10111 Sequence 2, Appli 1 1 1773.5 39.0 81 14 US-10-156-761-10111 Sequence 10111, A 1117 27.7 781 12 US-10-1395-241-14 Sequence 11, Appli 1 1117 27.7 812 12 US-10-1395-241-12 Sequence 11, Appli 1 1117 27.7 812 12 US-10-1395-241-12 Sequence 11, Appli 1 1117 27.7 812 12 US-10-1395-241-12 Sequence 18, Appli 1 111 27.7 812 12 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-13 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812 US-10-395-241-14 Sequence 18, Appli 1 111 27.7 812

Seguence 47; Appl		Sequence 4, Appli	Sequence 47453, A	Seguence 13955, A	Seguence 44566, A	Seguence 14105, A	Sequence 50616, A	Sequence 44567, A	Sequence 12, Appl	Sequence 11181, A	Seguence 44901, A	Sequence 8, Appli	Sequence 8, Appli	Sequence 9225, Ap	Seguence 64364, A	Sequence 63, Appl	Seguence 47312, A	Sequence 62548, A	Sequence 4, Appli	Sequence 4, Appli	Sequence 4, Appli	Sequence 4, Appli	Sequence 4, Appli	Sequence 4, Appli	Sequence 14161, A	Sequence 44953, A	Sequence 5176, Ap		Sequence 17890, A
US-09-927-827-47	US-10-282-122A-66335	US-10-246-330-4	US-10-282-122A-47453	US-10-369-493-13955	US-10-282-122A-44566	US-10-369-493-14105	US-10-282-122A-50616	US-10-282-122A-44567	US-10-245-802-12	US-10-369-493-11181	US-10-282-122A-44901	US-09-833-435A-8	US-10-375-720-8	US-10-156-761-9225	US-10-282-122A-64364	US-10-120-801-63	US-10-282-122A-47312	US-10-282-122A-62548	US-10-027-806-4	US-10-034-623-4	US-10-027-801-4	US-10-029-120-4	US-09-833-435A-4	US-10-375-720-4	US-10-156-761-14161	US-10-282-122A-44953	US-10-369-493-5176	US-10-369-493-5177	US-10-369-493-17890
10	12	14	12	15	12	15	12	12	14	15	12	σ	14	14	12	15	12	12	13	13	14	14	σ	14	14	12	15	15	15
555	2468	2468	2435	1465	1308	783	1074	1439	1119	613	1289	774	774	350	2204	2117	827	1721	3472	3472	3472	3472	176	176	1291	1325	1797	1805	1082
20.8	5.9	5.9	4.9	4.8	4.5	4.3	4.2	4.2	4.2	4.1	4.1	3.9	3.9	3.9	3.9	9.0	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.7	3.7	3.7	3.7	3.7
841.5	238	238	199.5	193	181	172.5	170	168.5	167.5	165.5	163.5	158.5	158.5	157	156.5	156	152	1:52	152	152	152	152	151.5	151.5	151	151	151	151	150.5
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

```
120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NMPGRCMGERLAVDPNNDNILYPGAPSGKGLWRSTDSGATWSQMTNPPDVGTYIANPTDT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 ATTOPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                             APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL TODERAL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197,4USOR
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 3
LENGTH: 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Score 4036; DB 12; Length 740; 100.0%; Pred. No. 2.2e-301; tive 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; OTHER INFORMATION: Catalytic domain GH74
US-09-917-376-3
                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                  Sequence 3, Application US/09917376 Publication No. US20040038334A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 100.
Matches 740; Conservative
US-09-917-376-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

180

120

240

300 3.00 360 420 420 480 480 540 540

```
AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
     NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                       NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                   TGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                        241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                  SGLIIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                   361 PWLTFGVQPNPPVPSPKLGMMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                  421 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                  NMPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDT
                                                                                                                                                                                                 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                 PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; OTHER INFORMATION: Catalytic domain GH74
US-10-155-400-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  721 YIGINGRGIVYGDIGGAPSG
   6
                                 61
                                                                                                                                  181
                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                661
                                                                  121
                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE:
                                                                  ò
                                                                                        g
                                                                                                                             ò
                                                                                                                                                            ద
                                                                                                                                                                                              ò
                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
 180
                                 240
                                                                  240
                                                                                                300
                                                                                                                                 300
                                                                                                                                                                 360
                                                                                                                                                                                                360
                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                 480
                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                                540
                                                                                                                                                                                                                                                                                                                                                                                                                                009
                                                                                                                                                                                                                                                                                                                                                                                                                                                               600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                421 APMYKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QPVAAGLPSSGAVGVMFHAVPGKEGDLMLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
NMPGRGMGERLAVD PNNDNILY FGA PSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                              SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                  SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                481 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                   KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                  241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                              PWLIFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                             PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                               APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                              AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                              WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA1TGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 6, Application US/09917376
Fublication No. US20040038334A1
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIWMEL, MICHAEL E.
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197-44501
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT APPLICATION NUMBER: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 740
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 ATTOPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                 TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Score 4036; DB 12; Length 740; llarity 100.0%; Pred. No. 2.2e-301; Conservative 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; ORGANISM: Acidothermus cellulolyticus
US-09-917-376-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 740; Conserv
                                 181
                                                              181
                                                                                                                                                               301
                                                                                                                                                                                              301
                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                               421
                                                                                                                                                                                                                                                                                                                                                                                                                              541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               721
                                                                                                                                                                                                                                                                                                                                                               481
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                            8
                                                      g
                                                                                           ò
                                                                                                                          qq
                                                                                                                                                              ð
                                                                                                                                                                                         g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

099

999

900 900 720

601 QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV 661 KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV Sequence 3, Application US/10155400

Sequence 3, Application US/10155400

Publication No. US20030108988A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: UNZANT, TODD B.

APPLICANT: HIMMEL, MICHAEL B.

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: NREL 01-36A

CURRENT FILING DATE: 202-10-22

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin Ver: 2.1

SEQ ID NO 3

LENGTH: 740 ö

9

9

```
NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                              NNWGYNGVYSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                      SGLTIDRQHDWTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QPVAAGLPSSGAVGVMFHAVPGKEGDLMLAASSGLYHSTNGGSSWSALTGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGFTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                  TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                    IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                            241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                   1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                   TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APMVKGLEETAVNDL1SPPSGAPL1SALGDLGGFTHADVTAVPST1FTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING; SHI-YOU
; APPLICANT: VINZARY, TODD B.
; APPLICANT: HIMMEL, MICHAEL B.
; APPLICANT: HIMMEL, MICHAEL B.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
                                                                                                                      ;
                                                                                          Length 740;
                                                                                     100.0%; Score 4036; DB 14;
100.0%; Pred. No. 2.2e-301;
ive 0; Mismatches 0;
                             TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VIGTNGRGIVYGDIGGAPSG
                                                                                                        Best Local Similarity 100.
Matches 740; Conservative
                                                                                                      Similarity
                                                ; ORGANISM: AC
US-10-155-400-6
SEQ ID NO 6
LENGTH: 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 721
                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                              241
                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            721
                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                          유
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                              ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                유장
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                     ò
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                               ö
                                                                                                                                                                                                                                                                                                                                   300
                                                                                                                      120
                                                                                                                                                  120
                                                                                                                                                                                 180
                                                                                                                                                                                                             180
                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           720
                                                                                     9
                                                                          ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                      I PHKGVFDPVNHVLY I ATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIMDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                           PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRWLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APWVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKWWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSA1TGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                       NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                  NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                 NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                            TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                       IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                               SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                           PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AELNPSI IVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                               ô
   740;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FROM ACIDOTHERMUS
Score 4036; DB 14; Length
Pred. No. 2.2e-301;
0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HOMEL, WICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM
TITLE OF INVENTION: CELLULLOLYTICUS
TITLE OF INVENTION: CELLULLOLYTICUS
TITLE OF INVENTION NUMBER: US/10/155,400
CURRENT APPLICATION NUMBER: 2002-10-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 6, Application US/10155400
Publication No. US20030108988A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YIGTNGRGIVYGDIGGAPSG
100.0%;
ilarity 100.0%;
Conservative 0
                 Similarity
Query Match
Best Local Simil
Matches 740; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-155-400-6
                                                                                                                       61
                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      721
                                                                                                                                                                                 121
                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                   g
                                                                                                                                              g
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                         욧. 명
                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              . <del>업</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                    ò
                                                                                                                                                                               8
```

•

540 540 600

480

99

660 720 720

600

300

360 360 420 420

180

180

120

240 240 300

```
TYPE: PRT
ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                          ; LOCATION: (957); OTHER INFORMATION: Any amino acid US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 100.
Matches 740; Conservative
                                                                                                                                                                                                                                                                                                            FEATURE:
NAME/KEY: MOD RES
                                                                                                                                                                                                                                                                 LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                            106
                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                        166
                                                                                                                                                                                                                                                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              994
                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                              ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              467 APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           527 AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                                                                                         NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDOGATWOITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                          107 NNWGYNGVVSIAADPINITNKVWAAVGMYINSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                   NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                      167 NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           287 IPHKGVFDPVMHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGOIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      541 WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         587 WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV
                                                                                                                                                                                                                                                                                                            1 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                100.0%; Score 4036; DB 12; Length 957; 100.0%; Pred. No. 3.2e-301; ive 0; Mismatches 0; Indels 0;
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.40801
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN VEY: 2.1
LENGTH. 957
                                                                                                                       TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         721 YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; OTHER INFORMATION: Any amino acid US-09-917-376-1
                                                                                                                                                                                                                                                Query Match 100.
Best Local Similarity 100.
Matches 740; Conservative
                                                                                                                                                     FEATURE:
NAME/KEY: MOD RES
LOCATION: (957)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           227
                                                                                                                                                                                                                                                                                                                                         47
                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
```

g

ò

g

ò

g

ð qq ò g

8

g

엄

ð

ò

엄

à

셤

a ò

ò

g ò g ò

Š

^

ö 360 420 480 540 706 120 107 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG 166 180 226 240 286 300 346 406 466 526 646 99 47 ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW 106 586 WAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF 600 167 NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGF 287 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY SGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE SGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAE AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV AELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV 1 ATTOPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGW NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT TGYOSDIOGVVWVAFDKSSSLGQASKT1FVGVADPNNPVFWSRDGGATWQAVPGAPTGF 241 IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGY PWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI APMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY APMVKGLEETAVNDL1SPPSGAPL1SALGDLGGFTHADVTAVPST1FTSPVFTTGTSVDY QPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG Gaps US-10-155-400-1

(Sequence 1, Application US/10155400

(Sequence 1, Application US/10155400

(Sequence 1, Application No. US200301089881

(GENERAL INFORMATION:

(APPLICANT: DING, SHI-YOU

(APPLICANT: DING, SHI-YOU

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: HIMMEL, MICHAEL B.

(APPLICANT: TELLORYTION: TELLUCLYTICUS

(APPLICANT: TELLORYTION: TELLUCLYTICUS

(APPLICANT: TELLING DATE: 2002-10-22

(APPLICANT: MICHAEL BATER: 2002-10-22

(APPLICANT: PARCHILING DATE: 2002-10-22

(APPLICANT: PARCHILING DATE: 2002-10-22

(APPLICANT: APPLICANT: 00.0%; Score 4036; DB 14; Length 957; 100.0%; Pred. No. 3.2e-301; ive 0; Mismatches 0; Indels 0; 647

ઠે 셤 ð 5 C

```
17;
                                                                                                                          658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISAEPWLTFGVQPNP-PVPSPKLGMMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS 414
                                                                              60 ---WINWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPF 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          177 PIDITGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGA 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        235 PQYGFLPHKGVLSPEEKTLYISYANGAGPYDGTNGTVHKYNITSGVWTDISP---TSLAS 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DYFCYSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         292 TYYGYGGLSVDLQVPGTLMVAALNCWWPDELIFRSTDSGATWSPIWEWNGYPSINYYSY 351
                                                           604
                                                                                                                                                                                          FGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLR 718
PSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFVWAPG 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A-GLPSSGAVGVMFHAVPGKEGDLWLAASS-----GLYHSTNGGSSWSAITGVSSAVNVG
                                                           545 DPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTFQPVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 ATTOPYTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: DING, SH1-YOU
APPLICANT: DING, SH1-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMEL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.4US01
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
41.6%; Score 1680; DB 12;
Best Local Similarity 46.2%; Pred. No. 2.7e-120;
Matches 346; Conservative 113; Mismatches 258;
                                                                                                                                                                                                                                                                                     747 RVYVSTNGRGIVYGDTAGSSDG 768
                                                                                                                                                                                                                                                                                                                                                                ; Sequence 7, Application US/09917376; Publication No. US20040038334A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                         RVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: PRT
; ORGANISM: Aspergillus aculeatus
US-09-917-376-7
                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-7
                                                                                                                          605
                                                                                                                                                                                                                                                         719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        296
485
                                                                                                                                                                                          629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 7
                                                                                    g
                                                                                                                                                       g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                            요
                                                           δ
                                                                                                                    ઠે
                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99 TGVVSLASDSVDPNKVYAAVGTYTNSWDPGNGAVLRSGDRGASWQKTDLPFKLGGNWPGR 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAITGDHANLRRV 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66 NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              396 GANPAPPEQTPKLGWMTESLEIDPFDSARMMYCTGATVYGTDNLTNWDSGSQFTIKPMAR 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTI 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISAEPWLTF 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVK 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLEETAVNDLISPPS-GAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELN 484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DIOGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFIPHKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39 YSWKNARVDGGGFVPGIVFNRSEKNLAYARTDIGGAYRWAESSKTWTPLLDSVGWSDWGH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 60.2%; Score 2429.5; DB 14; Length 882; Best Local Similarity 59.0%; Pred. No. 8.7e-178; Matches 438; Conservative 111; Mismatches 174; Indels 19;
                                                                                                                                                                                 740
                                                                                             ORGANISM: Streptomyces avermitilis
                                                                            YIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-156-761-9395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 9395
LENGTH: 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           219
                661
                                            707
                                                                            721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      156
```

유

ð 용 ò 임 ò 셤

ઠે

a ઠે 셤

ઠે

셤

ò

ሯ

 \mathcal{C}

φ

g ò g 8 g

à g ∂ qq

ò 임

à

```
Sequence 2, Application US/10420191
; Sequence 2, Application US/10420191
; Publication No. US20040067569A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040067569Alozymes Biotech, Inc.
; APPLICANT: Rey, Michael W.
; APPLICANT: Zaretsky Elizabeth J.
; APPLICANT: Basa, Jeffrey A.
; TILE OF INVENTION: Polypeptides Having Xyloglucanase Activity And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; TITLE OF INVENTION: Encoding Same
; FILE REFRENCE: 10210.200-US
; CURRENT FILING DATE: 2003-04-18
; PRIOR PRILING DATE: 2003-04-18
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Version 3.2
; SEQ ID NOS: 6
; SOFTWARE: Patentin Version 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20;
                                                                                                                                                                                                                                                                                           468
                                                                                                                                                                                                                                                                                                                                                                                                          523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                           408
                                                                                                                                                                                                                                     474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DGDTVLLMSSTSGALVSKSOG---TLTAVSSLPSGAVIASDKSDNTVFYGGSAGAIYVSK 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          196
      DYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVL 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79
                                                                                                                                                                                                                                                                                                                                                                                        292 TYYGYGGLSVDLQVPGTLMVAALNCWWPDELIFRSTDSGATWSPIWEWNGYPSINYYYSY
                                                                                                                  DISAEPWLTFGVQPNP-PVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                    GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 NNWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 YTWSNVAI-GGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWV----GW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . 44 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 838;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 45.00, 122... Mismatches 264;
Matches 325; Conservative 123; Mismatches 264;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39.1%; Score 1579; DB 12; 43.0%; Pred. No. 1.9e-112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        712 GDHANLRRVYIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GDLQTYGRVFRGHERPGHLLRQSQREPAG 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: Trichoderma reesei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-10-420-191-2
      296
                                                                                                                                                                           352
                                                                                                                                                                                                                                     415
                                                                                                                                                                                                                                                                                              409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          137
                                                                                                                     356
                                                                                                                                                                                                                                                                                                                                                    475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                         g
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                   g
                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ଚ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      17;
                                                                                                                  468
                                                                                                                                                                           534
                                                                                                                                                                                                                                     523
                                                                                                                                                                                                                                                                                           594
                                                                                                                                                                                                                                                                                                                                                    580
                                                                                                                                                                                                                                                                                                                                                                                                          653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : | | | | : | | | : | | | | : | | | | : | | | | : | | | | : | | | : | | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 PIDTIGYOSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGA 236
DISNAPWIQDTTSTDQFPV---RVGWMVEALAIDPFDSNHWLYGTGLTVYGGHDLTNWDS 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIAN 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PT-GFIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTAN 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59
                                                                                                                                                                                                         654 AVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQA--IT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60 NDTWHDW---GIDALATDPVDTDRVYVAVGWYTNEWDFNVGSILRSTDQGDTWTETKLPF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PQYGFLPHKGVLSPEEKTLYISYANGAGPYDGTNGTVHKYNITSGVWTDISP---TSLAS
                                                                                       DGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFYRST
                                                                                                                                                                                                                                                                                                                                                                                                       DGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI - TGVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---WINWGYNGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDOGATWOITPLPF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SSST--YTSDPVGIAWVTFDSTSGSSGSATPRIFVGVADAGKSVFKSEDAGATWAWVSGE
                                                      GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTT
                                                                                                                                                                     475 GTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 ATTQPYTWSNVAI-GGGGFVDGIVPNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: DINEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, WICHAEL E.
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: NREL 01-36A
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATCHIN VEY. 2.1
SEQ ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 14; Length 726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
41.6%; Score 1680; DB 14;
Best Local Similarity 46.2%; Pred. No. 2.7e-120;
Matches 346; Conservative 113; Mismatches 258;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GDLOTYGRVFRGHERPGHLLROSOREPAG 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDHANLRRVYIGTNGRGIVYGDIGGAPSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 7, Application US/10155400 Publication No. US20030108988A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Aspergillus aculeatus
US-10-155-400-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 9
US-10-155-400-7
                                                                                                                                                                                                                                                                                                                                                                                                       595
                                                                                                                                                                                                                                                                                        535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117
                                                         415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

ద

ð

g

ò

g

ò

g

∂

ò

ઠે ద ò 셤 ò g õ 셤 8

```
415 RHNVSIQSLADGIEBEFSVQDLASAPGGSELLAAVGDDNGFTFASRNDLGTSPQTVWATPT 474
  61 HNW---GIDAVALDPQDDQKVYAAVGMYTNSWDPSNGAIIRSSDRGATWSFTNLPFKVGG 117
                                                                         118 NMPGRGAGERLAVDPANSNIIYFGARSGNGLWKSTDGGVTFSKVSSFTATGTYIPDPSDS 177
                                                                                                                                           181 TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADP-NNPVFWSRDGGATWQAVPGAPTG 239
                                                                                                                                                                                                                                         FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFG 299
                                                                                                                                                                                                                                                               300 YSGLTIDROHPNTIMVATOISWWPDTIIPRSTDGGATWTRIWDWTSYPNRSLRYVLDISA 359
                                                                                                                                                                                                                                                                                                                                                                 586 VSKDTGSSF---TRG-PKLGSAGTIRDIAAHPTTAGTLYVSTDVGIFRSTDSGTTFGQVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA--YRSDDCGTTWVLINDDQHQYGNW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NMPGRGMGERLAVDPNNDN1LYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTY1ANPTDT
                                                                                                                                                                                                                                                                                                                                                                                                                                 360 EPWLTFG-VQPNPPVPS----PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                            415 GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHA---DVTAVPSTIFTSPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FITGISVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      532 ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               592 RSTDGGVTFQPVAAGLPSSGAVGVM--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           707 GOAITGDHANLRRVYIGTNGRGIVY--GDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          598 STKVAGSGSTAGQVYVGTNGRGVFYAQGTVGGGTGG 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: OWURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: BHIRAWA, JUN
APPLICANT: SHIRAWA, HIROSHI
APPLICANT: SHIRAW, HIROSHI
APPLICANT: SHIRA, TADAYOSHI
APPLICANT: HATTORI, WOSHIVKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REPERENCE: 249-226
CURRENT PILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: UP 2001-204089
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-06-30
PRIOR PILING DATE: 2001-06-30
PRIOR PILING DATE: 2001-06-30
PRIOR PILING DATE: 2001-06-30
SEQ ID NO 10111
LENGTH: 739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 10111, Application US/10156761; Publication No. US20030119018A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ) ORGANISM: Streptomyces avermitilis US-10-156-761-10111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-10-156-761-10111
                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           472
                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                           <del>Q</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                 ઠ
                                                                                             유
                                                                                                                                         ò
                                                                                                                                                                                유
                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGYQSDIQGVVWVAFDKSSSSLGQASKTIFVGVADP-NNPVFWSRDGGATWQAVPGAPTG 239
                                                                                                                          314 FGGLGLDLQKPGTLVVASLNSWWPDAQLFRSTDSGTTWSPIWAWASYPTETYYYSISTPK 373
                                                                                                                                                                                                                                                                                                                  374 APWIKNNFIDVTSESPSDGLIKKLGWMIESLEIDPTDSNHWLYGTGMTIFGGHDLTNWDT 433
                                                                                                                                                                                                                                                                                                                                                                                                           34 RHNVSIQSLADGIEEFSVQDLASAPGGSELLAAVGDDNGFTFASRNDLGTSPQTVWATPT 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          549 YSADGDILIWSTASSG---VQRSQFQSFASVSSLPAGAVIASDKKTNSVFYAGSGSTFY 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RSTDGGVTFQPVAAGLPSSGAVGVM--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI- 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     606 VSKDTGSSF---TRG-PKLGSAGTIRDIAAHPTTAGTLYVSTDVGIPRSTDSGTTFGQVS 661
                      YSGLTIDROHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRYVLDISA 359
                                                                                                                                                                                                                                                                                       360 EPWLTFG-VQPNPPVPS----PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDS 414
                                                                                                                                                                                                                                                                                                                                                                                     GGQIHIAPMVKGLEETAVNDLISPPSGAPLISALGDLGGFTHA---DVTAVPSTIFTSPV 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     649 TGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGA--YRSDDCGTTWVLINDDQHQYGNW 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 662 TALTNTYQIALGVGS-GSNW-NLYAFGT--GPSGARLYASGDSGASWTDIQGSQGFGSID 717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60
                                                                                             240 FIPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ASADGSRFVWAPGDPGQPVVYAVGFGNSWAASQGVPANAQIRSDRVNPKTFYALSNGTFY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 YTWSNVAI-GGGGFVDGIVFNEGAPGILYVRIDIGGMYRWDAANGRWIPLLDWV----GW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 39.0%; Score 1573.5; DB 14; Length 818; Best Local Similarity 42.9%; Pred. No. 4.9e-112; Matches 324; Conservative 122; Mismatches 265; Indel8 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Goedegebuur, Frits
APPLICANT: Goedegebuur, Frits
APPLICANT: Ward, Michael
APPLICANT: Ward, Michael
APPLICANT: Ward, Michael
TITLE OF INVENTION: EGVI Endoglucanase and Nucleic Acids
TITLE OF INVENTION: Encoding the Same
FILE REFERENCE: GG-994
CURRENT APPLICATION NUMBER: US/10/026,994
CURRENT FILING DATE: 2002-04-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GQAITGDHANLRRVYIGTNGRGIVY -- GDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STKVAGSGSTAGQVYVGTNGRGVFYAQGTVGGGTGG 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Application US/10026994 Publication No. US20030113732A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-10-026-994-2
                                                 197
                                                                                                                                                                                          300
                                                                                                                                                                                                                                                                                                                                                                                     415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               707
```

471

9

a

ò

셤

8

유

8 셤

ò

g

8

ò

414

531 528 591

31;

120;

124

63

241

```
219 TGAFFDKKPASIAPQPMKVALTP--NFLYVTYADYPGPWGVTFGEVWRQNRTSGAWDDIT 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | |: ||||| :: :|||||||::||: |: |
WWMSAVLIDPFNPEHLMYGTGATIWATDTLSRVEKDW------APSWYLQIDGIEENA 446
                                                                                                                                                                                                                                                                                     RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTPVPWLDFNNGPQWGGYGAPHGTPGLTKFG 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTT---GGTVAASADGSRFVWAP--G 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPOPVAAGLPSSGAVGVMFHAVPGKEGDLWL-AASSGLYHSTNGGSSWSAI-TGVSSAVN 656
                                                                                                                                                                                                                                             64 MCTESIALDPNNPDRLYLAQGRYVGDEW----AAFYVSEDRGQSFTIYESPFPMGANDMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRVGNSSPAPYNNQTFPAGGFCGLSVDATNPNRLVVIT-LDRDPGPALDSIYLSTDAGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NGVVSIAADPINTNKVWAAVGMYT-NSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                  SDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI---
                                                                                                                                                                                                                                                                                                                                                                                                         ------PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----SPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTI---IFRSTDGGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WTRIWDWTS------YPNRSLRYVLDISAEPWLTFGVQP-----NPPVPSP---KLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WMDEAMAIDPFNSDRMLYGTGATLYATNDLTK----WDSGGQIHIAP----MVKGLEETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSSGHEYDSACARGAYATDGGDAWTIPPTCPPGMNASHYQGSTIAVDASGSQIVWSTKLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DPGQPVVYAVGFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALSNGTFYRSTDGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VGFGKS----APGSSYPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 789;
                                                                                                                                      Indels
                                                                                                       27.7%; Score 1117; DB 12; 35.0%; Pred. No. 5.8e-77;
                                                                                                                                    281; Conservative 124; Mismatches 279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TGDHANLRRVYIGTNGRGIVYGDI 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EADPKVYGRVYLGTNGRGIVYADL 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-10-395-241-12
; Sequence 12, Application US/10395241
; Publication No. US20040038367A1
; GRURRAL INFORMATION:
SOFTWARE: PatentIn version 3.2 SEQ ID NO 14
LENGTH: 789
TYPE: PRT
ORGANISM: Geotrichum sp. M128
                                                                                                                      Similarity
                                                                           US-10-395-241-14
                                                                                                                                                                                                                          99
                                                                                                                                                                                                                                                                                     125
                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   680
                                                                                                                                                                                                                                                                                                                                              185
                                                                                                                                                                                                                                                                                                                                                                                                         242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                711
                                                                                                        Query Match
                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 14, Application US/10395241
Publication No. US20040038367A1
GENERAL INFORMATION:
APPLICANT: YAOI, KATGULO
APPLICANT: MITSUISHI, Yasushi
TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
FILE REFERENCE: 073756
CURRENT APPLICATION NUMBER: US/10/395,241
PRIOR APPLICATION NUMBER: US 2002-03-35
PRIOR PILING DATE: 2002-03-25
NUMBER OF SEQ ID NOS: 19
                                            21,
                                                                                                     101
                                                                                                                                    125
                                                                                                                                                                160
                                                                                                                                                                                             185
                                                                                                                                                                                                                          207
                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                   257
                                                                                                                                                                                                                                                                                                                 293
                                                                                                                                                                                                                                                                                                                                            317
                                                                                                                                                                                                                                                                                                                                                                                                       368
                                                                                                                                                                                                                                                                                                                                                                                                                                    413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ANDYFGYSGLTIDRQHPNTIMVATQISWWPDTIIFRSTDGGATWTRIWDWTSYPNRSLRY 353
                                                                         65
                                                                                       SAKVPLRAAYDTHTRELYVTYGDAPGPGGOSDGSVHKLRTATGTWTEVTPVKPGGTTSDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----APRIRGLEESAVRQLISPPVGEAHLISGLGDIGVMYHERLTASPSRGMATNPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DIQGVVWVAFDKSSSLGQASKTIFVGVADPNNP----VFWSRDGGATWQAVPGAPTGF
                                                                                                                                                                                                                                                                                                              ---IPHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRISPV----PSTDT
                                                                                                                                                                                                                                                                                                                                                                                         SADTFAYGGVAVDARRPGTLVVSTNNRWADGDTVFRSTDGGRTWTSLKD-----AA
                                                                                                                                                                                                                                                                                                                                                                                                                                 VLDISAEPWLTFGVQPNPPVPSPKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FTTGTSVDYAELNPSIIVRAGSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTTGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NGVVSIAADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPGR
                                                                                                                                                                                            GMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               VPDVSETPFLDWGDD-----KPKFGWWIQALAVDPYDSQHVVYGTGATLYGTRDLKRW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGGQIHIAPMVKGLEETAVNDLISPPSG-APLISALGDLGGFTHADVTAVPST-IFTSPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VAASADGSRFVWA----PGDPGQPVVYA----VGFGNSWAASQGVPANAQIRSDRVNPKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FYA--LSNGTFYRSTDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAAS-SGLYHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TNGGSSWSAITGVSSAVNVGFGKSAPGSSYPAVFVVGTIGGVTGAYRSDDCGTTWVLIND
                                                                        6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                              Gaps
                                            86;
                Length
               Query Match 34.8%; Score 1405; DB 14; Best Local Similarity 40.8%; Pred. No. 3.8e-99; Matches 308; Conservative 104; Mismatches 257;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DQHQYGNWGQAITGDHANLRRVYIGTNGRGIVYGD 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DAHOWGWIGEAVVGDPRIHGRVYLATNGRGIQYGE 737
                                                                                                                                 99
                                                                                                                                                                                          126
                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                     186
                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                           258
                                                                                                                                                                                                                                                                                                                                                                        294
                                                                                                                                                                                                                                                                                                                                                                                                    318
                                                                                                                                                                                                                                                                                                                                                                                                                                 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           703
                                                                                              셤
                                                                                                                                                        g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                               ò
                                                                                                                                                                                       ⋧
                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

335

379

431

336

491

505

565 598 621

710

738

```
GENERAL INFORMATION:

APPLICANT: Yasushi
APPLICANT: Yasushi
APPLICANT: MITSULSHI, Yasushi
TITLE OF INVENTION: HOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
TITLE OF INVENTION: BNCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
TITLE OF INVENTION: BNCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
TITLE REFERENCE: 073756
CURRENT APPLICATION NUMBER: US/10/395,241
CURRENT FILING DATE: 2003-03-25
PRIOR APPLICATION NUMBER: JP 2002-83433
PRIOR PLING DATE: 2002-03-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PATENTIN VERBION 3.2
SEQ ID NO 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Histidine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S-----VIFDP-----ERNGTIYASATAPQG-MYVTHDGGVSWEPVAGQPSSWLNRT 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65 MGTESIALDPUNPDRLYLAQGRYVGDEW----AAFYVSEDRGQSFTIYESFFPMGANDMG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           220 TGAFPDKKPASIAPQPMKVALTP--NFLYVTYADYPGPWGVTFGKVWRQNRTSGAWDDII 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             337 WKDVTQLSSPSNI,EGNWGHPTNAARY-KDGTPVPWLDFNNGPQWGGYGAPHGTPGLTKFG 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTT---GGTVAASADGSRFVWAP--G 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             432 VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRA 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 NGVVSIAADPINTNKVWAAVGMYT-NSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----SPVPSTDTANDYFGYSGLTIDROHPNTIMVATQISWWPDTI---IFRSTDGGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            337 WTRIWDWTS-----YPNRSLRYVLDISAEPWLTFGVQP----NPPVPSP---KLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    380 WMDEAMAIDPFNSDRMLYGTGATLYATNDLTK----WDSGGQIHIAP----MVKGLEETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with US-10-395-241-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels 120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 826;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 27.6%; Score 1114; DB 12; Best Local Similarity 35.0%; Pred. No. 1e-76; Matches 281; Conservative 123; Mismatches 280;
                 762 EADPKVYGRVYLGTNGRGIVYADL 785
711 TGDHANLRRVYIGTNGRGIVYGDI
                                                                                                                               Sequence 18, Application US/10395241 Publication No. US20040038367A1
                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                             US-10-395-241-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
ઠે
                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠ
         APPLICANT: MISULEHI, Yasushi
TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME,
TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
FILE REFERENCE: 073756
CURRENT APPLICATION NUMBER: US/10/395,241
CURRENT FILING DATE: 2003-03-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PATENTIN VERSION 3.2
SOFTWARE: PATENTIN VERSION 3.2
LENGTH: 812
                                                                                                                                                                                                                                                                                                                                                              31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            242 IGAFPDKKPASIAPQPMKVALTP--NFLYVTYADYPGPWGVTFGEVWRQNRTSGAWDDIT 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              599 TPQPVAAGLPSSGAVGVMFHAVPGKEGDLWL-AASSGLYHSTNGGSSWSAI-TGVSSAVN 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SYAAKGAGLUT - - GTSLMPAVNPWVAGDVWVPVPEGGLFHSTDFGASFTRVGTANATLVS 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 RGMGERLAVDPNNDNILYFGAPSGKGLWRSTDSGATWSOMTNFPDVGTYIANPTDTTGYQ 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185 SDIQGVVWVAFDKSSSSLGQASKTIFVGVADPNNPVFWSRDGGATWQAVPGAPTGFI--- 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   359 WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTPVPWLDFNNGPQWGGYGAPHGTPGLTKFG 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   432 VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDYAELNPSIIVRA 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSFDPSSQPNDRHVAFSTDGGKNW--FQGSEPGGVTT---GGTVAASADGSRFVWAP--G 544
                                                                                                                                                                                                                                                                                                                                                                                                   65
                                                                                                                                                                                                                                                                                                                                                                                                                                       86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      300 PRVGNSSPAPYNNQTFPAGGFCGLSVDATNPNRLVVIT-LDRDPGPALDSIYLSTDAGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 MGTESIALDPNNPDRLYLAQGRYVGDEW----AAFYVSEDRGQSFTIYESPFPMGANDMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------PHKGVFDPVNHVLYIATSNTGGPYDGSSGDVWKFSVTSGTWTRI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WTRIWDWTS-----YPNRSLRYVLDISAEPWLTFGVQP-----NPPVPSP---KLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         380 WMDEAMAIDPFNSDRMLYGTGATLYATNDLTK----WDSGGOIHIAP----MVKGLEETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VGFGKS----APGSSYPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 YTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPLLDWVGWNNWGY
                                                                                                                                                                                                                                                                                                                                                                                                                         | : ||||||| : ||| : || : XEFKUVAIGGGGYITGIVAHPKTXDLLYARTDIGGAYRWDAGTSKWIPLNDFIEAQDMNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NGVVSIAADPINTNKVWAAVGMYT-NSWDPNDGAILRSSDQGATWQITPLPFKLGGNMPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----SPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQISWWPDTI---IFRSTDGGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      545 DPGQPVVYAVGFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALSNGTFYRSTDGGV
                                                                                                                                                                                                                                                                                                                                                            Indels 120;
                                                                                                                                                                                                                                                                                                                         Length 812;
                                                                                                                                                                                                                                                                                                                       Query Match 27.7%; Score 1117; DB 12; Best Local Similarity 35.0%; Pred. No. 6e-77; Matches 281; Conservative 124; Mismatches 279;
                                                                                                                                                                                                                                               ; TYPE: PRT
; ORGANISM: Geotrichum sp. M128
US-10-395-241-12
APPLICANT: YAOI, Katsuro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                703
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            유
                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
        QY
        545 DPGQPVVYANGFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALSNGTFYRSTDGGV 598

        Db
        567 EQASGPWYSHDYGKTWS----VPAGDLKAQTANVLSDRVQDGTFYATDGGKFPVSTDGGK 622

        QY
        599 TFQPVAAGLPSSGAVGVWFHAVPGKEGDLML-AASSGLYHSTNGGSSWSAI-TGVSSAVN 656

        Db
        623 SYAAKGAGLVT--GTSLMPAVNPWVAGDWVPVPEGGLFHSTDFGASFTRVGTANATLVS 680

        QY
        657 VGFGKS----APGSSYPAVFVVGT--IGGVTGAYRSDDCGTTWVLINDDQHQYGNWGQAI 710

        Db
        681 VGAPKSKSDGKKASAPRAVFWGTDKPCSDIGLYRSDDNGSTWTRVNDQEHNYSG-PTMI 739

        QY
        711 TGDHANIRRYIGTDYGGTVYADI 734

        Db
        740 EADPKVYGRVYLGTNGRGIVYADI 763
```

Search completed: May 11, 2004, 12:25:47 Job time: 90.3468 secs

Run on:

```
Pseudomon
Bordetell
                                                                                                                                                                                                                                                                                                                                                                                                                            Aspergill
Aspergill
                                                                                                                                                                                                                                                                                                                                                                                   Clostridi
Clostridi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Oryza sat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Xanthomon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Thermotog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Xanthomon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Geotrichu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unculture
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Vibrio pa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Burkholde
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Vibrio pa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pseudomon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Geobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Methanosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pseudomon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rhizobium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gloeobact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus muscu
                                                                                                                                                                                                                                                                                                                                                                    Caldicel]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bordetel]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gloeobact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bordetel]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BX842573 Mycobacte
D90904 Synechocyst
D90915 Synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hypocrea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bordetel
                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                Agaricus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Agaricus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AJ133614 Bacillus
                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AK110506
AP005031
AP005031
AE011712
AE012276
AX924431
AX924433
AX924433
AX924435
AX924435
AX924435
AX924435
AX924435
AX92435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AY328853 AP006570 AY357582 AY328852 AE004613 AE017211 AE011079 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF182518 | AE016774 | BX640426 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AL603645
BX640440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BX640425
AP006578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AL671876
BX640439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AL445199 1
AP005221
AF178428 1
                                                                                                                                                                                                                                                                                                                                                                                                                                          AY040839
AY281371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                           AX700058
AP005028
SCO939128
AX565635
AF078038S1
AE007608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABI534359
AY28135752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AX924431
AX924429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RME603645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                \L445199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AP006578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AY328853
                                                                                                                                                       em_htgo_other:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
                                                                                                              em_sy:*
em_htgo_hum:*
                                                     em_htg_pln:*
em_htg_rod:*
                                                                                  em_htg_mam:*
                                                                                                 vrt:*
                                           em_htg_mus:*
                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                    Query
Match Length
                                                                                                                                                                                                                                                                                                             2869
299800
296500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47537
301950
58128
49575
20342
300134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   299550
14520
13278
11910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        306050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        348866
                                                                                                                                                                                                                                                                                                                           2429.5
2371
2063
2061.5
2009
1982
1753.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1579
1560.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1405
1162.5
1160.5
                                                                                                                                                                                                                                                                                                              4036
                                                                                                                                                                                                                                                                                                                                                                                                                                          1620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                226.5 218.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   204.5
198.5
193.5
193.
192.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        192
189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  460.9
368.9
271.9
268.9
                                                                                                                                                                                                                                                                      Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                       Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-MODEL=frame+ p2n.model -DEV=xlh
-MODEL=frame+ p2n.model -DEV=xlh
-MODEL=frame+ p2n.model -DEV=xlh
-G=Capt_1/GSPTO spool/USO9917376/runat_11052004_114535_28241/app_query.fasta_1.1429
-DB=GenEmbl -QFWT=fastap -SUPFIX=p2n.rge -MIRMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=buman40.cdi -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TAINGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIZE=560 -MINLEN=0 -MAXIEN=200000000
-USER=USO9917376_@CGN 1 1_7742_@runat_11052004_114535_28241 -NCFU=6 -ICFU=3
-NO MAAP -LARGEQUERY -NEG SCORES=0 -WART -DSPBELOK=100 -LONGLOG
-DBV TIMEOUT=120 -WARN -NEG SCORES=0 -THERADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                   (without alignments)
3015.368 Million cell updates/sec
                                                                                                                                                                   .....YIGTNGRGIVYGDIGGAPSG
                                                                                  May 12, 2004, 03:40:05 ; Search time 10636.8 Seconds
                                                                                                                                                                                                                                                                                                                6940544
 GenCore version 5.1.6 (c) 1993 - 2004 Compugen Ltd.
                                                       OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                   3470272 segs, 21671516995 residues
                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                             Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                    1 ATTOPYTWSNVAIGGGFVD
                                                                                                                                                                                                                                                                                                                                          seq length: 0
seq length: 200000000
                                                                                                                                        US-09-917-376-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gb_un: *
gb_vi: *
em_ba: *
em_fun: *
em_fun: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        _8t8:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_or:*
em_ov:*
em_oh:*
em_ol:*
em_ol:*
em_ro:*
em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gb ba:*
gb_htg:*
gb_om:*
gb_ov:*
gb_bat:*
gb_p1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_in:*
em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   *:EO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gb_sy:
                                                                                                                                                                                                 BLOSUM62
                                                                                                                                           Title:
Perfect score:
                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                          Minimum DB
Maximum DB
                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database
```

		381 MethapgluhlamethlaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 400	441 GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 460	1639 AACGACAGGCACGTCCCGACAGACGGCGCAAACTCCAAGGCAGGAA 1698 521 ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 540 1699 CCTGGCGGGGTGACGACGGCGCACCGTCGCCGACGACCGAC
8 8 8 8 8 8 8 8 8	8 6 8 6 8 6	8 8 8 8 8	8 8 8 8 8 8	4 6 8 6 8 6 ·
LOCUS LOCUS LOCUS AX700058 AX700058 ACCESSION AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:29536021 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX700058.1 G1:2953602 AX7	2869 stive: 740 nes: 0 0 0 0 lAlalleGlyGlyGlyGlyPheValAsp	139 GCGACGACTCAGCCCTACACCTGGAGCAACGTGGCGATCGGGGGGGG	ValtrpalaalaValQlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaileLeu GTATGGGCCGCCGTCGGATGTACACCAACAGCTGGGACCCAACGACGAGCGATTCTC ArgSerSerAspGlnGlyAlaThrTrpGln11eThrProLeuProPheLysLeuGlyGly ArgCatCGTCTGATCAGGGCGAACGTGGCAAATAACGCCCTGCCGTTCAAGCTTGGCGGC AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAshAspAsnIle	16.1 TrpSerGlnWetThrAsnPheProAspValGlyThrTyrlleAlaAsnProThrAspThr 16.1 TrpSerGlnWetThrAsnPheProAspValGlyThrTyrlleAlaAsnProThrAspThr 16.1 TrpSerGlnWetThrAsnPheProAspValGlyThrTyrlleAlaAsnProThrAspThr 16.1 TrpSerGlnWetThrAsnPheProAspValGlyThrTyrlleAlaAsnProThrAspThr 16.1 ThrGlyTyrGlnSerAspIleGlnGlyValValThrPyalAlaPheAspLysSerSerSerSerSerThrGlyTyrGlnSerAspIleGlnGlyValValValTrpValAlaPheAspLysSerSerSerSerSerThl

us-09-917-376-3.p2n.rge

임 ð g ò 名 ò 셤 ઠ 셤 ઠે 셤 Š 셤 ઠ a ò

```
/translation="waggrainavpecadidtipmtsaesdgpgrvdrilrypvksmlgee
i.pavevterglahprrvalidrdfkvasakhpreimarlitytaqadnsaventsbg
kelmstdpsvdbealseivsrpvkitappperaaldradpedvlrdhtrpppvsartgri
gggspegtfpdpapythymtstlaalaalbydsaqaeryrpnivlturtpltrgeten
widrtihigtglvlrviartprealptipmtstrans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /trānslation="MGQPTPARIQKDWRTAVIYQVYVRSFADGDGDGTGDLAGVRAR
LPYLAGLGYDAIWFNPWYLSPMKDGGYDVADYRAVDPAFGTLAEAEKLIAEARELGIR
             Tel:81-3-3481-1933, Pars 18-3-3481-8424)
This work was done in collaboration with Harto Ikeda(*1), Jun Enkikawa (*2), Aktharu Hanamoto(*3), Chiquas Takahashi(*3), Mayumi Shinose(*3), Hiroshi Horikawa (*4), Hidekazu Nakazawa (*4), Tomomi Osonoe(*4), Norihiro Kushida(*4), Hisashi Kikuchi(*4), Tadayoshi Shiba(*5), Yoshiyuki Sakaki(*6,*7), Masahira Hattori(*1,*7) and Satoshi Omura(*1,*3).
Final finishing process and all annotation were done by H. Ikeda and J Ishikawa
                                                                                                                                                                                                                                          *! Kitasato University
*2 National Institute for Life Sciences, Kitasato University
*2 National Institute of Infectious Diseases
*3 The Kitasato Institute
*4 National Institute of Technology and Evaluation
*5 School of Science, Kitasato University
*6 Institute of Medical Science, University of Tokyo
*7 RIKEN, Genomic Sciences Center
Following url is also available.
http://avermilis.ls.kitasato-u.ac.jp.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Streptomyces avermitilis MA-4680"
/mol type="genomic DNA"
/strain="MA-4680"
/db xref="taxon:27882"
/note="This strain is also named as strain: ATCC 31267,
NCIMB 12804 or NRRL 8165.~synonym: Streptomyces
(B-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATGNTATVTLPSGTNLRYLKIANTANTGWPAAQFSEVEAYLS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product="putative alpha-glucosidase"
/protein_id="BAC69435.1"
/db_xref="G1:29605369"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCAGVYAQVLRPGRIRLGDPVHWA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /transT_table=11
/product="hypothetical_p:
/protein_id="BAC69433.1"
/db_xref="GI:29605367"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /gene="aglA1"
complement(3464. .5065)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complement (1232. .3406)
/note="SAV1723"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement (3464. .5065)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /gene="aglA1"
/note="SAV1724"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="SAV1722"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /codon_start=1
/transl_table=:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           avermectinius'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                  COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S.
Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis
Nat. Biotechnol. 21 (5), 526-531 (2003)
                                                                                                                                                                                                          1998
                                                                                                                                                                                                                                                                                                                                                                                                                     2118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2179 ACGGGGCGTACCGCTCCGACGACTGTGGGACCTGGGTACTGATCATGATGATGACAG 2238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2239 CACCAATACGGAAATTGGGGACAAGCAATCACCGGTGACCACGCGGAATTTACGGCGGGTG 2298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2299 TACATAGGCACGAACGCCCGTGGAATTGTATACGGGGACATTGGTGGTGCGCCGTCCGGA 2358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BCT 10-MAY-2003
  1819 GCTGCTTCGCAAGGTGTTCCCGCCAATGCCCAGATCCGCTCAGACCGGGTGAATCCAAAG 1878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   721 TyrlleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                             640
                                                                                                                                                                                                                                                                                                                                                                  99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (19-MAR-2002) Director-General of Biotechnology Center, National Institute of Technology and Bvaluation, Biotechnology Center; 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
                                                  ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAgpGlyGlyValThrPhe
                                                                                                        1939 CAACCGGTCGCGGCCGGTCTTCCGAGCGGTGCCGTCGGTGTCATGTTCCACGCGGTG
                                                                                                                                                                                                                                                             ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn
                                                                                                                                                                                                                                                                                                                1999 CCTGGAAAAGAAGGCGATCTGTGCTCGCTGCATCGAGCGGGCTTTACCACTCAACCAAT
                                                                                                                                                                                                                                                                                                                                                                                                                     2059 GGGGGCAGCAGTTGGTCTGCAATCACCGGCGTATCCTCCGCGGTGAACGTGGATTTGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2119 AAGTCTGCGCCCGGGTCGTCATACCCAGCCGTCTTTGTCGTCGGCACGATCGGAGGCGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgVal
                                                                                                                                                     GlnProvalAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Omura,S., Ikeda,H., Ishikawa,J., Hanamoto,A., Takahashi,C.,
Shinose,M., Takahashi,Y., Horikawa,H., Nakazawa,H., Osonoe,T.,
Kikuchi,H., Shiba,T., Sakaki,Y. and Hattori,M.
Genome sequence of an industrial microorganism Streptomyces
                                                                                                                                                                                                                                                                                                                                                                  GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Streptomyces avermitilis genomic DNA, complete genome, section
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 (bases 1 to 299800)
Omura,S., Ikeda,H., Ishikawa,J., Hanamoto,A., Takahashi,C.,
Shinose,M., Takahashi,Y., Horikawa,H., Nakazawa,H., Osonoe,T.,
Kushida,N., Director-General of Biotechnology Center, Shiba,T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Streptomyces avermitilis MA-4680
Streptomyces avermitilis MA-4680
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001) 21477403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   avermitilis: deducing the ability of producing secondary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             299800 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sakaki, Y. and Hattori, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AP005028.1 GI:29605366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AP005028 BA000030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   metabolites
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AP005028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11572948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12692562
                                                     581
                                                                                                                                                          601
                                                                                                                                                                                                                                                                621
                                                                                                                                                                                                                                                                                                                                                                     641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
AP005028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
```

LOCUS

g

```
AKDPDLPDLVEGRAPHPFVDRELHDIYRSWRAIADEYGAVFVGEVWLPDTERFARYL
RPDELHTAFNFSFLSCPWDARRLRTAIDETLAEHAPVGAPATWVLCNHDVTRTVTRYG
RTDTGPDFAAKARGTPFDLALGTRRARAAALISLALPGSVIYGYGEEGLPERAIPRE
SIQDPWHFRSHGVDPGRDGCVPLPWVADAPHAGFGGEPWLPQPARWSAYAADLQAED
PGSWLSLYREALRARAYADGPLTWLPAPDGVLAFARTEGLMCVVNLADTPAGLPAHS
RLLLSSGPLDDGGRLPQDTAAMLRI
COMPIEMENT (5097. .5972)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complement(5969..6925)
/note="SAV1726"
/note="SAV1726"
/codon garat=1
/transl_table=11
/product="putative integral membrane protein"
/protein id="BAG69437.1"
/db_xref="G1:25065341"
/translation="MTKTAARRPVEAIAVRPVPAPPPGGGGGRRRLADQAHAYAFLLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NEALERIHLPTSNWSNGADTALVSLVTVATWANWGGTTLIYLAALQGIPGELYERAEL
DGANILQRIRHVTIPQTRFIVLMLMLLQIIATWQVFTBPFVITGGGPENATVTVLXI
KYXPLYNDFGGACALSVMLLVLLGAFSAVYLRLTRSAEEDS"
COMplement (6931. .8271)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /codon start=1
/transl_table=11
/product="putative sugar-binding lipoprotein precursor"
/protein id="BACG6438.1"
/db xref="G1:2565372"
/translation="MRSSGFRRACVALVACSFVLTACGSSGDAADGKTRITVNCMPPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SAKVDRKFFEDDIAAFEKKNPDIDVVAHDAFPCQDPKTFDAKLAGGQMEDVFTTYFTD ATHVDVNQAADITSVYKDLKSYDTVQQLRDIYTVDGKYVGYPRTGSEMGLIYNRGL FERACLDPDREPATBEVRAAAKRI BALGGGTVGSADSSAQNGGGWHFTAELYSQGG VASADGKKATVDFPEGKAVLQNLHDMWTDNSMGSKQLIVINDVQQLMGSGKLGMYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APDNIPILVKEKGGTYTDLALAPMPGGKGTLIGGDGYMFNKKATPAQIKAGLKWLDFM
TLTGKGRELGDYARAKKNDAPVGLPEPRLFSGAADARDOQVKKANANVPVENYQSFLD
GNOSLRWKIEPPQAQIYSVLDSAVSAVLTKKDADIDKLLKDASGKIDSILARG"
complement (8418. 9410)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /transl_table=11
/product="putative LacI-family transcriptional regulator"
/protein_id="BAC69439.1"
/db xref="GI:28665373"
/translation="MTRRLAEVAKKVGVSEATVSRVLNDKPGVSPATRQAVLSALDVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GYERPTQLRGERARLVGLVLPELQNPIPPAFAEVIGGALAQLGLTPVLCTQTKGGVSE
ADYVNLLLQQQVSGVVPAGGLYAQADAPHDHYRRLADRNIPVLVNAAIEHLGPPAVS
CDDAVAVEQAMRHLASLGHER IGLVLGRGHMPSARKLAARALGDLPDEYVARARFS
IEGGHAAASRLIDRGYTGFICASDPLALGAYRAARRKGLDVPSQISVVGYDDSALMNC
TEPPLTTVRQPIESMGRAVVELLINARIGGSSVPDEBLLIFEPELLVRGSTAQAPR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /trānslation="wsgdpapvpgdpaprrxtamaaqhspalptlggrkpsyrervad
Arraaliagergerysapglaaregysatpyreamidlerselydagelydyrervy
Vsrgldbythirslielpftaglaltadpaelealrpvareivaargblayvea
DlefthiggllalagophivDvyrdlrrrsrlygitalaragrlaasaseellellohlla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /tränslation="MSTRTLVSPLTLSRPRGKAAYWTVFTGVVLLFALAFLFPVYWWV
TARMKSDPBARTPPSLV9GXMLLSGYTDAWDLWALLWHWTVWQATGAWLLQUVFC
TAAAYALSRLKFVFGGTLVLGGTLATLWVPRYLTVWDRYLTVADLPLIHTSLLNDPLGI
WLPAVANAFNLYLLKRFFDQIPRDVLEAAEIDGAGRLRTLWSIVLPMSSRVLGVVSIF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALVAVWQDFLWPLMVFSDTDKQPISVALVQLSQNIQLTVLIAAMVIASIPMVAMFLVF
QRHIIAGISAGSTKG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLLCFALFSWYPAIRAVVIAFQKYTPGSPPEWVGTANFTRVLHDPEFTÄAWRNTLFT
LLALLIGFAVPFVMALVLNELRHAKAPPRVVVYLPVMIPPUVSALLWKWFYDPGAGLA
TLVDI VPNHVSDQHPWFRAALAAGPGSPERALFHFRGGRGAHGELPPNDWPSQFAGST
EPVWTQLPDGDWYLHLFTPEQPDLNWAHPAVRQEHEDVLRFWFERGVAGVRI DSAALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GntR-family transcriptional regulator"
                                                                                                                                                                                                                                                                                                                                                                                                                                                       membrane protein"
                                                                                                                                                                                                                                                                                                                                                                                                         /transT_table=11
/product="putative integral
/protein_id="BAC69436.1"
/db_xref="G1:29605370"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'product="putative GntR-1
'protein id="BAC69440.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        complement (9457. .10185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                db_xref="GI:29605374"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        table=11
                                                                                                                                                                                                                                                                                                                                       'note="SAV1725"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'note="SAV1728"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'note="SAV1729"
                                                                                                                                                                                                                                                                                                                                                                              codon start=1
trans table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon_start=1
/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 codon start=1/transl table=
```

CDS

CDS

CDS

CDS

```
175817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      176117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      176237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /trānslation="MRSKLVLHAVDSHTEGMPTRVITGGIGTIPGATMNERRLYFREH
TSDDTRQLLMNEPRGHSAMSGAVLQPPARPDCDVGVVYIEVSGYLPMCGHGTIGVATVL
VETGMVEVVEPVTTIRLDTPAGLVVAEVAVEDGAARAVTLKNVPSFCVGLDRKATLAD
GRTVTYDLAYGGNPFALILPLGEFGLEPDRSRKDDILKAGLSLMDAVNALGGPVHPEDP
SIRGCHHVHLVAPGSDARHSRHAMAIHPGWPDRSPCGTGTSARMAQLHARGELPLGTE
                                                                                                                                                                                 FVNESLIGTRFTGRLLGTTEVAGIPAVLPSFTGRAWYTGTAQYLLDPEDPFPAGFVL"
complement (11214. .12104)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlalleLeuArgSerSerAspGln 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhelleProHisLysGly 245
                                                                                                                                                                                                                                                                                                                                                                                                25
                                                                                                                                                                                                                                                                                                                                                                                                                                                     45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                                                                                                                                                                  175938 ACGGGTGTCGTGAGCCTCGCCTCCGACCGAACCAAGGTGTACGCGGCCGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                     GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaalaasnGlyargTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlyAlaThrTrpGlnIleThrProLeuProPhelysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CCCAGGGGCAAGGGGTTGTGGGCGGTCGACGGACTCGGGGGGCCTCCTGGTCGCAGGTCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AspileGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                        racemase
                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                     proline
RDTEAAREVTTRHLGHVRGLWVAP"
complement (10198. .11199)
/note="SAV1730"
                                                                                                                                                                                                             /gene="dapA1"
complement(11214. .12104)
                                                                                  /protein_id="BAC69441.1"
/db_xref="GI:29605375"
                                                                                                                                                                                                                                                                                                                                                                   (1-299800)
                                           /codon_start=1
/transl_table=11
/product="putative ]
                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-3 (1-740) x AP005028
                                                                                                                                                                                                                                                                    1.25e-113
2429.50
73.99%
59.03%
60.20%
                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                        Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     226
                                                                                                                                                                                                                                                                                                                                                                                                                                                     56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               206
                                                                                                                                                                                                                                                                                                                                                                                               ø
                                                                                                                                                                                                                                                                                                                                                                                                                          175758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         176058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                176118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      176178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             176238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176298
                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                       Alignment
                                                                                                                                                                                                   gene
                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                               ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ጵ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

177468 TTCAAGTCCGGCAGGTTCTACGTCGACGGGGGGGGGCGACCTTCACGGCGTCCGCG 177527	Stropponyces coelicolor A1(2) complete genome; segment 25/29. Streptomyces Coelicolor A1(2) complete genome; segment 25/29. AL939128 AL033862 AL031013 AL031017 AL031350 AL031371 AL031515 AL033505 AL03443 AL035161 AL079355 AL066825 AL645882 AL939128.1 GI:2441307 Streptomyces coelicolor A3(2) Streptomyces coelicolor A3(2) Streptomyces coelicolor A3(2) Streptomyces coelicolor A3(2) Streptomyces coelicolor A3(2) Bacteria, Actinobacteria, Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces. Bentley,S.D., Chater,K.F., Cerdeno-Tarraga,A.M., Challis,G.L., Thomson,N.R., James,K.D., Harris,D.E., Quail,M.A., Kleser,H., Harper,D., Bateman,A., Brown,S., Chandra,G., Chen,C.W., Collins,M., Rabinowitench,E., Rajandream,M.A., Rutherford,G., Hornsby,T., Howarth,S., Rabinowitench,E., Rajandream,M.A., Rutherford,G., Grares,S., Taylor,K., Warren,T., Wietzorrek,A., Woodward,J., Barrell,B.G., Parkhill,J. and Hopwood,D.A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2) Nature 417 (6885), 141-147 (2002) 2 (bases 1 to 296500) Bentley,S.D. Submitted (09-MX-2002) Submitted on behalf of the Streptomyces sequencing team, sanger Institute, Welloome Campus, Hinxton, Cambridge CB10 183 B-mail: sdbesanger.ac.uk On or before Oct 26, 2002 this sequence version replaced g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520757, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520757, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520757, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520757, g1:20520759, g1:20520757, g1:20520757, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:20520759, g1:
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	ACCOSTOR SCORE COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR COCOSTOR
	QY 406 ThrhambspleuthrbysTrpaspSerGlydJvGlnIleHisIleAlaProMetValLys 425 Db 176949 AcGGACACCCGAACCCGAACCCGATCACCCCCCCCCCCCC

gene

Sec

CDS

```
ARHR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGS
                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref="SPTREMBL:09ZBJ4"
/translation="MTYGCIASMKTRPGFREEVVSLLLSAADGLRDAGCELYVVGLSE
DDETTIWVTEVWRTREDHDASLELPAAKDAIGRAMPMLTGEFTKQEMSVAGGLGL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /trānslation="MAQVEÄTTERIVAADAETVPDTLADYSGTREKLLPEHFSEYER
EGGDGEGTLVHWKLQATSKRVRDCLLEVTEPTDGELVEKDRNSSMVTTWRVTPAGEGK
SRVVVLTTWQGAGGVGGFFEKTFAPKGLARIYDALLAKLAAEVEK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db_xref="SPTREMBL:09ZBJ3"
/translation="MRWENLTDSDHARAADTALFGADAVTRTFDTPEFRGITFHEVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARSILINRVPGASRMPFEWTVNPYRGCTHACVYCFARKTHSYLDLDTGLGFDSQIVVKV
NAPEVLRRQLASRRWQGEHVAMGTNVDCYQRAEGRYRLMPGIIEALTERANPFSITK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTLILRDLDLLTRAARVTEVGISVSVGFTDAELWRTVEPGTPAPERRLEVVRALGERG
IGCGVLMAPVIPFLGDEPAQLRATVRAIAAAGATSVTPLVLHLRPGAREWFMAWLGRH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HPHLVRRYERLYADGAYAPKWYQRRITRQVHDLAREYGIGPARAGMPRRVAEPEHSRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="SC9C7.18, conserved hypothetical protein, len: 348 aa; similar to two hypothetical proteins e.g. TR:069685 (EMBL:AL022121) Mycobacterium tuberculosis RV03718c or YRV025.066c, 147 aa; fasta scores; opt: 475, z-score: 768.3, E(): 0, (52.1% identity in 146 aa overlap)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="SC9C7.17c, conserved hypothetical protein, len: 348aa; similar to several hypothetical proteins eg. SW:YOBN_MYCTU from Mycobacterium tuberculosis (340 aa) fasta scores; opt: 958, z-score: 1321.8, E(): 0, (47.6$ identity in 355 aa overlap)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note="SC9C7.19, probable efflux protein, len: 815aa; similar to many, several of which are involved in toxin /antibiotic efflux eg. TR:069070 (EMBL:AF061335) oxytetracycline exporter from Streptomyces rimcsus (536 aa) fasta scores; opt: 1020, z-score: 879.9, E(): 0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAA22730.1"
/db_xref="G1:4154077"
/db_xref="SPTREMF1..007077"
                                                                                                                                                                                                                                                                                                                    /product="hypothetical protein SC9C7.16c"
/protein_id="CAA22728.1"
/db_xref="GI:4154075"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAA22729.1"
organism="Streptomyces coelicolor A3(2)"
                                                                                                                                                                                                                                              force="SCOCT.16c, unknown, len: 99 aa"
(codon start=1)
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (36.9% identity in 542 aa overlap)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement (517. .1563)
/gene="SCO6481"
/note="synonym: SC9C7.17c"
complement (517. .1563)
/gene="SCO6481"
                                                                                                                                                                     'note="synonym: SC9C7.16c"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1800. .2243
/gene="SCO6482"
/note="synonym: SC9C7.18"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'note="synonym: SC9C7.19"
                                                                                 xref="taxon:100226"
                          'mol_type="genomic DNA"
/strain="A3(2)"
                                                                                                           complement (144. .443)
/gene="SCO6480"
                                                                                                                                                                                            complement (144. .443)
/gene="SCO6480"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1800. .2243
/gene="SCO6482"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2394. .4841
/gene="SCO6483"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2394. .4841
/gene="SCO6483"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AEPAPPEPVQLSLI"
```

gene

CDS

gene

CDS

```
/translation="MGGTILVQDPVRDPVQDEPVVAAQAPPGPPVPGPDGERDAGLSP
RRYRLVFVGLMLALLIAALEQMIVATLPKIVGELHGLDKNSWATTAYLLTGSTVGLPI
YGKLGDLFGRKGVFQFAILVFVVGSALAGRAQTMDQLIAFRAVGCVGAGGLMIGVQAI
IADIVPPRERARYNGLIGASFGLASVAGPLLGGYFTDHLSNRWCFYINVPFGIYTWAV
VAVVILKLPRPRVKPRLDILGALLLAAASTCLVLLTSWGGTEYAWGSRVILGLAAGAVA
                                                                                                                                                                                                                                                                                                                ATLLFLVAEHFAPEPLJ PLRLFRDSI FNVTALVGLVVGVALFGAASYLPTFLQMVDGA
SATESGLLIMPMGGVVGASI ISGQLI SRTGHTRIPLIGSALSVVGWALLSRLDADT
SRLHYS IWMAVLGAGIGLVMPVLVLAVQNYRPTDLGTATGANNY FRQI GGSVGAAVF
GTLFAGRLIPDALADR I PARAGVGLPDARAITPQLVHSLPPALKDAYI GAYADAMPR I F
                                                                                                                                                                                                                                                                                                                                                                                                                                          LYLVPVLVUGLLIALPLKEKPLVSHNAPVTDPBTAPHAPQPQPAAQVPPARSPYAGGI
PVCGTVQHPDGTVVPRAALTLIDVGCQQIGRGASGDDGRYALATPGSGAYVLIAAAGG
HQPQAVSVTYGERPVELDVYGGAGRLAGSVMTABGSPVRDATVTLTNVHGEVVATTR
SGREGGYVITELVASAPPRPAALPVSVQAREFRQDVELAGGAVLKGTVR
ASGGRAVEDARVTLLDAAGNVDTLTTGPDGTPRFVDLSSGEYTVIAAGGVPVATVU
VAGGGRTERDLQLGHED"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="conserved hypothetical protein"
/protein id="CAA22732.1"
/protein id="CAA22732.1"
/db_xref="di-144079"
/db_xref="coa:0928J0"
/db_xref="spTREMBL:0928J0"
/translation="WYGHPRTVVTATQILRSRGQRGPDAMDHGTERDAPPGHRAGPD
TVPADPVTGRVPLAVVVVDROGLVSHWSRGARRLFALSKEBAIGRPAPDLLPVSGVLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PFPAHTPFFQCMTTGEPVLVPRITEEMGHAIASQFEKRDIRPLITGRSMLLVPLKART
VVLGFMILLRHPERPVFNDMDRVTGAELAARAGLVLDNARMYTHQENVAETLQDSMLP
HIPARLAGCDIATRYLPGTLLGRVGGDWFDAVKLPGARTALVVGDVMGHGLNSAAMMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AYPLVGPGPERLLVLAADADGLRRDYADGTDDSAGPAVVERLAPAFALHTDFPGADEL
ARRLPEILPSMSVGESARIVAQVLELGYPVLEFSQHDRVPVTPDWGVPRRAERRARRE
RAARALAAGEPVPDELRDEAEDLEYAAVRERLEFLNEVSGAIGTSLDLSRTIVEVSRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QLRTAVQTMAGLDLPPAQLLRNLDDLAQRLGDTHLATCLYAVYDPIAGELHLANAGHI
PPVLVRAEDGASELIDLPTGAPIGVGGVPFESVRVTVAPQBDRLVMCTDGLVEVRGEDI
GVGLATLCESAAHPAASMDDACDTIIRALGAGGGRKDDVALLMARLTGIEPDAVAEWR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VVPRFTDVAGTYLREQVVAGEGFPDGVPDTTTMMHRVALEHTDEPGRMDDVVPVGEAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LARDPAEVGRARAAVREQLYDWGLPKLVAPAELMVSELVTNAVRHSHARPVALRLVRA
DTLLCEVDDDDHELPALKSAGPGDETGRGLRVVSTLAREWGASRTGAGKTVWFELTLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EDAEPDGAGHDDSYRSYTAHDGPGPGLESSLDGRLSYPAAGRARLTVRADDRVDVLWW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /product="hypothetical protein SC9C7.21"
/protein id="CAA22733.1"
/bx.xref="G1:4154080"
/db.xref="G0A:092B19"
/db_xref="G0A:092B19"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7711. 8433
/gene="SCO6485"
/note="SC9C7.21, unknown, len: 240 aa"
/product="putative efflux protein"
/protein id="CAA22731.1"
/db_xref="G1:4154078"
/db_xref="GOA:092801"
/db_xref="SPTREMBL:092801"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note="synonym: SC9C7.20"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  note="synonym: SC9C7.21"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /gene="SCO6484"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'gene="SC06485"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon_start=1
/transl_table=:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4911. .7493
```

```
SerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeulleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArg
                                                                                                                                                     GGCACGGAGAACCTGACGAACTGGGACGACGACGGCGGCGCCACGTTCGCCGTCGAGCCGATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyr
                                                                                                                                                                                                                                                                                                                                                   ThrilellePheArgSerThrAspGlyGlyAlaThrTrpThrArglleTrpAspTrpThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                    PheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetAlaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerThrilePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----aaggirgccggcgrgcagracaccggcrrcggracgrccrggcaggcgrc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyPro
                                                                                                                                                                                                       IleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr
                                                                                                                                                                                                                                                                               305 IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Accescroccescescescesarcerceastes and a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a second a 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AlaThrAsnAspleuThrLysTrp---AspSerGlyGlyGlnIleHisIleAlaProMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       265
                                                                                                                                                                                                         285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         69368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      69128
                    69779
                                                        245
                                                                                           69719
                                                                                                                                                                     69629
                                                                                                                                                                                                                                             69599
                                                                                                                                                                                                                                                                                                                   69548
                                                                                                                                                                                                                                                                                                                                                       325
                                                                                                                                                                                                                                                                                                                                                                                           69488
                                                                                                                                                                                                                                                                                                                                                                                                                               345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 69308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                89069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      69008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              68963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     68903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             68843
                                                                                                                                                         g
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  જે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        중 음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                         કે
                                                                               윱
                                                                                                                   ર્જ
                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                     $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             70080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    70320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   70200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       70020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              69840
/translation="MSVTSRYRDAWEGFWNEASGDRGAVFWDAEPTLTAGRHLAHFEP
HTTDAALPLIDLGCGNGTQTRYFABRFPHVVGADLSAAALAHARHADPGGLAAYRQLD
AADKAEAQLAHAELGDANVVVGGYLHQCEPDDRQPLADALAALLGERGRAFLVBPSQA
AKYVLMGLAQCPAGPAKLAPVLRHGLTPGAVPDEAVPEYLRAAGLTVLASGELPLVT
TEYTADGTRIELPSAWVVAGTA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85
                                                                                                                             /note="synonym: SC9C7.22"
8603. .9436
/gene="SC06486"
/note="SC9C7.22, possible transport associated protein, len: 277aa; similar to SW:DPPA BACSU dppA, dipeptide transport protein (274 aa) fasEa scores; opt: 448, z-score; 604.3, E(): 2.3e-26, (32.3% identity in 269 aa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GAGGAGTCCCACACCCTGGACGCCGCTCCTCGACCACGTCGGCTGGGACGACTGGGGGGCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TyrThrTrpSerAsnValAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGCATGGGTGAGCGCCTGGCCGTCGACCACGACGACGTGTACCTGGGCGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CCCAGGGGCCACGGACTGTGGGCGGTCCACCGACGCGGCGTCACCTGGTCCGAGGTGACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AspileGinGlyValValTrpValAlaPheAspLysSerSer---SerSerLeuGlyGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                   296500
431
109
183
24
10
                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (1-296500)
                                                                                                                                                                                                                                                                               /codon_start=1
/transl_table=11
                                                                                            8603. .9436
/gene="SCO6486"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    x SC0939128
                                                                                                                                                                                                                                                                                                                                                     1.24e-110
2371.00
72.29%
57.70%
58.75%
                                                                                                                                                                                                                                                              overlap)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740)
                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                      Alignment Scores
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    70379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         70199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       70079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  205
                                                                                                                                                                                                                                                                                                                                                                                                            Best Local S
Query Match:
                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           $ B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
```

404

423

443

463

344

284

68904

68844

563

603

```
KVSGVQEGDAVGFGKATSSSGYPVIFTSAKIDGVRGIFRSDDEGKTWKRINDNHNQWY
WTGASITCDPDVEGRVYIGTNGRGIIVGDSSTPPPFGQDSGSASGWTGATPTPETPDS
PHTGGSTDTGETSEPGTGSGSGTAPDSGNTSPGNTSSCKVRYSTTDWGSGFTGAATI
TNTSPPSPINGWTLAFTYPSGQTISSWWSATQTLSGRNVVLKNSGWNPTIPAGGSVEFG
FNGAYTSKNTPPTSFSLNGVTCALAD"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     265 TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArg 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       130 TATAGCTGGTCGAACGTTGAGATTGTGGCCGCCGTTTTGTGCCTGGAATTGTCTTTAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    66 AsnGlyValValSerIleAlaAlaAspProlleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      310 TCCGGAATTCTGTCCTTGGCCACTGACCCTGTTGACACCCGCGTGTACCTTGCAGCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGly---TyrGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerAspileGinGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGGCAAAACCAAGGTGTTGTGTGGGTAACTTTTGACCCCACATCAGCGAAAGCTGGTCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            205 AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       245 GlyvalPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr
                                                                                                                                                                                                                                                                          2823
386
1112
225
18
                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                               (1-2823)
                                                                                                                             1. .69
70. .2820
/product="unnamed"
                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-3 (1-740) x AX565635
                                                                                                                                                                                                                                                          4.16e-97
2063.00
67.21%
52.09%
                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                             sig_peptide
mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           225
                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                 ..
0
                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          윰
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KKVDITVKAQGIBETAAQDLAAPPGNIGLYSALADIGGFTHKDISQVPNKYYYKNPHH
DAVYBIDFBEBEKPATVVRAGKBISGETTSWVQYSTDAGETWRCATPSGVKGRGSITV
SANASSIVAAPEGAPRRSTNSGSSWSTVSGLPHNAQVASDRVNANTLYGFVDGKFYH
STNGGASFTASAFTGFPTSGNVRFRAVPGRQCHLWLAGGVSGSTYGMWRSTDGGKWWT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /protein_id="CAD58415.1"
/db_xref="G1:26000970"
/db_xref="RI:26000970"
/db_xref="REMTRENED:CAD58415"
/tanslation="MRLRTVVATALTALATTAGGIAATSTPMPVTAPPSFSSSPSGY
SWSNVEIVGGGFVPGI-VFNQKDPGLVYARTDIGGAYELNDSTGRWIPLTDHIGMDDWS
HSGILSLATDPVDTNRVYLAAGTYSMDWDPQNGAILRSADKGETWEKTMLPFRVGGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YSGLAIDRKNPDTIMVVSQVSWWPDMYVYRSTDRGKTWSPIWELNGSQPRTKQYNHDY
SGAPWLDFGNTAKEPEANPKLGWMTQSFEIDPHNSDRFFYGTGAGIYGGTNLTNWDKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YTGQNQGVVWVTFDPTSAKAGQTTQTIYVGVADKQNNVYRSTDGGATWQRVPGQPTGF
LAQKGVLDHKGQQLYIATSDTGGPYDGSKGDVWRLDISSGQWTRISPIPSTSSNSAFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 29-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PGRGMGERLAVDPNNNKVLYFGAESGNGLWKSTDYGKTWGKVTSFPNAGNYVADASGA
                                                                                                                                                      637
                                                                                                                                                                                                                                                   657
                                                                                                                                                                                                                                                                                                                                                677
                                                                                                                                                                                                                                                                                                                                                                                                                                               697
       TCGACGGACGGCGGACCTTCACCAGCTGCCGGTGTCGACGCGGCGGACACCCGTC
                                                                                gescanceccencercie cresces con secon contra con contra con contra con contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra co
                                                                                                                                                                                                                                                                                                                                                GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrlle
                                                                                                                                                                                                                                                                                                                                                                              GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeu
                                                                                                                                                      LysGluGlyAspLeuTrpLeuAlaAlaSerSer-------GlyLeuTyrHis
                                                                                                                                                                                                                                                SerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gacdaceccaccadregederegaccedecedecarcacegedeacegederac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           718 ArgArgVallyrIleGlyThrAsnGlyArgGlyIleVallyrGlyAspIle-----
                                                       AlaAla---GlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear

    .2823
    /note="unnamed protein product"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="Jonesia sp., DSM 14140"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Family 74 xyloglucanases
Patent: WO 02077242-A 1 03-OCT-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1 from Patent WO02077242. AX565635.1 GI:26000969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    organism="unidentified"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GCCGGGGGCACCGATCCGGGT 68235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Novozymes A/S (DK)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---GlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1, .2823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Duffner, F. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unidentified
unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unclassified
                                                       604
                                                                                                                                                    623
                                                                                                                                                                                                   68615
                                                                                                                                                                                                                                                                                                                                                                                               68495
                                                                                                                                                                                                                                                                                                                                                                                                                                               678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             735
       68729
                                                                                                      69989
                                                                                                                                                                                                                                                     638
                                                                                                                                                                                                                                                                                                   68555
                                                                                                                                                                                                                                                                                                                                                  658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              68315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          68255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
AX565635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
       셤
                                                                                                    셤
                                                                                                                                                 ò
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                       8
```

45

309

65

189

105

429

369

125 489 145 549 165 609 184 999 204 720 224 840

244

900

697

2199

셤

셤 ò 셤 ઠે 셤 ò ద

ò

셤

ò

g ò 셤

ò

ð 셤 8

à

ò g g δ 셤 δ 셤 ઠે 셤 à

ò

ò

```
GERLAIDPNINRILYGGTREGXGLWKSEDYGVSWKKVTSFPNPGTYIEDPNCPNDYLN
HTGVVWWYDPDFSGRPGEXKIITYVROGADKTYSTYTKOGGOTWQALAEDEOPTGLED
RAKLSSOGHLYITYSTYNTGEPROLISPENDYGEVWRYNTKTGEWKOISPWAAQDTYFGYGGLAD
DAQNPKVVWVAALSSWWPDTYIWRSIDGGSTWKCIWEWNGYPNRTLHYNMDISAAPWL
                                                                                                                                                                                                                                                                                                                                                                                                                                2260 GGGCGCGTCTACATTGGTACCAACGGTCGCGGCATTATCGTTGGTGATTCCTCCACTCCC 2319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EGGGGFITGIVFNPKEKNLVYVRTDIGGAYRSTDGGYTWTQLMDWVSFDDMNLLGVES
IATDPVDPNXVILACQGTYTNSWTDMNGAILRSTDEGDTFEITPLPFXLGGNMPXRNL
  1960 CGTCAAGGTCATCTCTGGCTTGCTGGTGGAGTATCGGGAAGCACCTATGGGATGTGGGGG 2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BCT 11-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene, partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product="glycosyl hydrolase 5"
/protein id="AAK06388.1"
/db_xref="G1.1X743879"
/translation="11X78LISEFSILLILNFLIAEFLQIGAKADTGIAVQPYVWKNVKI
                                                                                                                                                                                                                                                                                                                   AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeu 717
                                                                                                                                                                                                                                                                                                                                                                                                      718 ArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAla 737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Macquarie University,
                                                                                                                                                               2140 GATGGAGTTCGAGGAATTTTCCGGTCTGACGATGAAGGCAAGACCTGGAAACGCATCAAC
                                                                                                                                                                                                                                                                                                                                          GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsn
                                                                                                                                   GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIle
                                         SerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Finciutes; Clostridia; Clostridiales; Syntrophomonadaceae; Clostridia; Clostridiales; Syntrophomonadaceae; Caldicellulosiruptor.

Gibbs, M.D., Reeves, R.A., Farrington, G.K., Anderson, P., Williams, D.P. and Bergquist, P.L.
Multidomain and multifunctional glycosyl hydrolases from textreme thermophile Caldicellulosiruptor isolate Tok7B.1 curr. Microbiol. 40 (5), 333-340 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         glycosyl hydrolase 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 (bases 1 to 3262)
Gibbs, M.D., Reeves, R.A., Farrington, G.K., Anderson, P. Williams, D.P. and Bergquist, P.L.
Direct Submission
Submitted (27-OCT-1999) Biological Sciences, Macquari
Sydney, NSW 2109, Australia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Caldicellulosiruptor sp. Tok7B.1"
/mol_type="genomic DNA"
/isolate="Tok7B.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref="taxon:80339"
274. .>3262
/note="ORF5; multidomain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF078038S1 3262 bp Caldicellulosiruptor sp. Tok7B.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Caldicellulosiruptor sp. Tok7B.1
Caldicellulosiruptor sp. Tok7B.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1. .3262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 codon_start=1
transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF078038
AF078038.1 GI:12743874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2320 CCG 2322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   738 Pro 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10706665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of 5
                                                                                          2020
                                                                                                                                                                                                                           678
                                                                                                                                                                                                                                                                                                                   869
                                                                                                                                                                                                                                                                                                                                                               2200
                                                                                                                                      658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF078038S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
SEGMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                             g
                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                           8
                                                                                       셤
                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||||||| :::::::
TTTGGCAATACAGCCAAAGCTAACCCCAAACTCGGGTGGATGACCCAGTCC 1257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           438 TACTCAGGGCTGGCAGACATCGGTGGCTTCACCCACAAAAAATCTCTCAGGGGCTGAAC 1497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----GAAGGTGCCGCTCCTCGACGTTCCACCAACAGTGGCTCCTCGTGGTCAACGGTG 1785
                                                                                            |||||||||||:::||||||||||
ATCAGCCCAATTCCCTCTACC---TCATCGAATTCAGCGTTCGGGTACTCCGGTTTAGCC 1017
                                                                                                                                                                                                                                                                            1078 ATGTACGTATATCGCTCCACCCACCGGGGGAAAACGTGGTCGCCTATCTGGGAACTGAAC 1137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1258 TTCGAGATTGACCCGCATAACTCCGACCGCTTCTTCTATGGCACCGGCGCGGCGATCTAT 1317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAAĞĞĞATCĞAAĞAAACCĞCAĞCCCAAĞACCTTĞCCGCTCCCCTGGAAACATTGGCCTA 1437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TGGGTTGGGGTTTCAACTGATGCTGGCGAAACGTGGAAACCTGGCGCCACGCCATCAGGG 1671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       622
                                                                                                                                                             324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                     325 ThrileilePheArgSerThrAspGlyGlyAlaThrTrpThrArgileTrpAspTrpThr 344
                                                                                                                                                                                                                                                                                                                                        PheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAla 384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       385 MetalaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyr 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               405 AlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetVal 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeu 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             445 IleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSer 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484 AsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArg 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            584 AlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProVal 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---AlaSerSerGlyLeuTyrHis 637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Thr --- IlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu 483
IleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr
                                                                                                                                                          IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 544 GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         564 GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ala---AlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                derrreacegegrrreccaccagrega----aacerrecrrrecereccereccereccesee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LysGluGlyAspLeuTrpLeuAla--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              623
                         901
                                                                                                                 961
                                                                                                                                                             305
                                                                                                                                                                                                         1018
                                                                                                                                                                                                                                                                                                                                                                                        1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1732
                                                                                                                                                                                                                                                                                                                                                                                                                                    365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1612
```

£.

```
NFGXTBPTPPEVSPKLGWMVGTLEIDPFNSDXMLYGTGAXLYGCDDLTNWDKGQNITI
KVKAIGIEETSVQALISPPVGPHLFSALGDIAGFRHEDLEKAPNWTYVQPNMGTTTDI
DFAELNPNFWVRVGNVDKQWNPNTNRIGFSYDGGKSWFQGNTEPQGTSEGGTVAAAAD
                                       GSAVVWAPKGAKVCYSTDNGNKWVECANVPSEAIVYSDRVNPNKFYAFKNGKFYISAD
KGKTFIESPAAGLPISGNFKTVPGIEGDIWLVGNNGMWHSTDGGYSFVKISGVEDAAS
IGFGKPAEGETYPAIYTYAKINGVRGIFRSDDCDKTWIRINDDKHQFGCANADITGDP
                                                                              615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThr 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPhe
                                                                                                                                                                                                                                                                                                                                                                                                             GlnProTyrThrTpSerAsnValAlaIle---GlyGlyGlyGlyPheValAspGlyIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysVal---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                       376 CAGCCATATGTGTGGAAGAATGTTAAAATTGAAGGTGGAGGAGGTTTTATCACTGGAATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 496 CGTAGTACAGACGGTGGCAATACATGGACACACACTTATGGACTGGGTTAGTTTTGATGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GATTATTTGAACCATATAACAGGTGTGGTCTTGGGTATTTTGATCCTACAAGTGGTAGG
                                                                                                                                                                         /evidence-experimental
2791. .3262
/note-"type IIb cellulose binding site"
2581. .2790
/note-"PT-linker"
                                                                                                                                                   274. .2580
/note="Region: endoglucanase domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <del>:</del>
                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                      YGENAKVTLYVDGVLVWGQEPGGA"
                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-3 (1-740) x AF078038S1 (1-3262)
                                                                                                                                                                                                                                                                                     5.85e-97
2061.50
65.62%
52.58%
51.08%
                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                    misc_feature
                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                     sc_feature
                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                              4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               162
                                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 . 염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
1383
                                                                                                                                                                                                                                                                                                                 1443
                                                                                                                                                                                                                                                                                                                                                                                                               1563
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1923
           ||||||| || :::||| || || ||||:::||| || GGATTGGCAGCTCTTAGCTCATGG 1323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGACCACATTTATTCAGTGCACTGGGAGATATAGCGGGATTCAGACATGAAGACTTGGAG 1743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||||||||||||||||:: ::||||||
GCTGAGTTAAAITCCAAATTTTAIGGTTCGIGTAGGGAAIGTTGATAAACAGTGGAATCCA 1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                                                                                        380
                                                                                                                                                                                                                                                                                                                                                                                        400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  519
                                                                                                                                           301
                                                                                                                                                                                            321
                                                                                                                                                                                                                                           TrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp ·341
                                                                                                                                                          TrpThrArgileSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSer
                                                                                                                                                                                                                                                                  TGGCCAGATACCTATATTTGGAGAGTATCGATGGTGGCGAAACGTGGAAATGTATTTGG
                                                                                                                                                                                                                                                                                                                                                                                                        ||||
| ATGGTTGGAACACTTGAAATAGACCCATTTAATTCAGATARAATGCTTTATGGAACAGGA
TrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle
                                                                                                                                                                                            GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp
                                                                                                                                                                                                                                                                                          342 AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro
                                                                                                                                                                                                                                                                                                      TrpLeuThrPheGly---ValGlnProAsnProProValProSerProLysLeuGlyTrp
                                                                                                                                                                                                                                                                                                                                                       TGGCTAAACTTTGGAAWTACTRATCCTACACCACTGAAGTAAGCCCAAAACTTGGCTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                       401 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                    421 AlabroMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGly---Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAggGlySerArgPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro
                                               ProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThr
                                                                                                                                                                                                                                                                                                                                                                                         WetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     461 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GTTTGGGCACCA----AAAGGAGCGAAAGTATGCTATTCTACAGATAACGGAAACAAA
                                                                                                                     1153
                                                                                                                                                                                                                                           322
                                                                                                                                                                                                                                                                                                                                                                                         381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 540
                                                                       1096
                                                                                                                                                                    1213
                                                                                                                                                                                                                   1264
                                                                                                                                                                                                                                                                   1324
                                                                                                                                                                                                                                                                                                                 1384
                                                                                                                                                                                                                                                                                                                                                                                                               1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1924
                                                                                              262
                                                                                                                                             282
                                                                                                                                                                                             302
                                                                                                                                                                                                                                                                                                                                         362
                                                                                                                                                                                                                                                                                                                                                                  1444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  260
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                 g
                                                             셤
                                                                                            ò
                                                                                                                   셤
                                                                                                                                            8
                                                                                                                                                             a
                                                                                                                                                                                          8
                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                  ठ
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                ਨੇ
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                         임
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                               8
```

```
/protein_id="AAK78899.1"

/db_xref="G1:15023824"

/db_xref="G1:15023824"

/kranslad.ton="MOSLKVDWRNGLLKINBAKQGLKSAVDKFSLKNKSRSIEIELKN

/kranslad.ndp.egeirkinvnidywykrfydydkkaskikTrSgytyrkkVGLinFGE

EVERVPVIGGIGVPIDAVESWYDRNKFYLKKTYTITLETGAIGVCIVTLPEAGITWKA

WDALSAAGGVMSLDNIGTAASAIYQRKVERKSEEDSSGYDLFQSIFGGAGWYSAKILG
product="Probably secreted sialidase; several ASP-boxes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TVVPSFFVWMGVDGTEEWHKANYNLKEDALMVAANYFSDLALKVLSKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               function in bacteria is unknown"/protein id="AAK78896.1"
/db_xref="G1:15023821"
                     and dockerin domain"
/protein_id="AAK78895.1"
/db_xref="GI:15023820"
                                                                                                                                                                                                                                                                                                                                                                                                                                       complement (2634. .2978)
/gene="CAC0920"
                                                                                                                                                                                                                                                                                                                                                                                                       complement (2634. .2978)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5056. .6348
/gene="CAC0923"
5056. .6348
/gene="CAC0923"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3104. .4252
/gene="CAC0921"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3104. .4252
/gene="CAC0921"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4594. .4863
/gene="CAC0922"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4594. .4863
/gene="CAC0922"
                                                                                                                                                                                                                                                                                                                                                                                                                        /gene="CAC0920"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon_start=1
/transl_table=:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon start=1
/transl table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /codon_start=1
/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Childress, D., Zeng, Q. and Smith, D.R.
Childress, D., Zeng, Q. and Smith, D.R.
Direct Submission
Submitted (24-JUL-2001) GTC Sequencing Center Production,
Finishing, and Bioinformatics teams, Genome Therapeutics Corp., 100
Beaver Street, Waltham, MA 02453-8443, USA
Location/Qualifiers
                                                                                                                                                                                                                                                                                2446 GACAAACATCTGGTTGTGCGAATGCAGACATAACTGGAGATCCACGCGTGTACGGT 2505
                                                                                                                                                                                                             2206 ACTGTTCCAGGTATAGAAGGTGATATCTGGCTAGTTGGTAATAATGGGATGTGGCACTCA 2265
                                                                                                                                                                                                                                                                                                                                                                AE007608 12732 bp DNA linear BCT 27-JUL-2001 Clostridium acetobutylicum ATCC824 section 96 of 356 of the
                                           2098 AATAAGTTTTÄTGCATTTAAGAATGGCAAATTTTATATAAGTGCTGATAAAGGAAAAACA 2157
                                                                                                                           2158 TTTATCGAAAGTCCTGCTGCTGGTTTACCAATATCAGGA------AATTTCAAA 2205
                                                                                                                                                                      638
580 LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThr 599
                                                                                  Phe---GlnProValAlaAlaGlyLeuProSerGsrGlyAlaValGlyValMetPheHis 618
                                                                                                                                                                                                                                                                                                                                        PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGly 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Molling, J. Breton, G., Omelchenko, M.V., Markarova, K.S., Zeng, Q., Gibson, J., Breton, G., Omelchenko, M.V., Markarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, B.V. and Smith, D.R. Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acerobutylicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Clostridium acetobutylicum
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
                                                                                                                                                                    619 AlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaAserSerGlyLeuTyrHisSer
                                                                                                                                                                                                                                                      639 ThrAsnGlyGlySerSerTrpSerAlaileThrGlyValSerSerAlaValAsnValGly
                                                                                                                                                                                                                                                                                                                                                                                                                        679 GlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             699 AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .12732
/organism="Clostridium acetobutylicum"
/mol type="genomic DNA"
/strain="ATCC 824"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /db_xref="ATCC:824"
/db_xref="taxon:1488"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Clostridium acetobutylicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="CAC0919"
73. .2592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="CAC0919"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AE007608.1 GI:15023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 complete genome.
AE007608 AE001437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11466286
                                                                                    009
                                                                                                                                                                                                                                                                                                                                          629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE
PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
AE007608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                         셤
                                                                                                                         셤
                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                               쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

IATORIO I VINITALE MODIE AL RECORD AND TRANSMINITE DE SONO TO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRANSMINITE DE SONO TRAN /translation="MPFINSTVTVKLDKEKESTLKVELGKIIELLPCKSEDWLMVGFK DNYPLYFKGEKKDKAAFIEVKIFGSADKASKNKLVSAISSLMENELSIPKDCIYITVD EVSNWAWNGSLF" /codon start=1 /transl_table=11 product="Uncharacterized small conserved protein, homolog of yijA/yukE B subtilis" /protein_id="AAK78898.1" /db_xref="GI:15023823" /product="Metal-dependent amidohydrolase" /protein id="AAK7889".1" /db_xref="G1023802" /translation="MITIEKQESFLKKLIGYRRELHKHPELSMKEYETTKRIKRWLYE NDIKVLDFPIEVGVVAEIEGEHEGKTVAIRADIDALPIEEKTELPFSSVNKGIMHACG HDFHTAAILGTAIILSKRKAEIYGRVRIIFQPGEETGKGSQYIIEEGVLRDVDCIFGM HNKPDLPVGTIGIRSGELMASVDRFEIKVIGSGGHAGIPNKCIDPIVVASQIVASMQT VVSRSISPIESVVISITRFNSGTTWNVISDSAEMEGTVRTFNNNIREEIRNLMERTAK GIARALGGKIEFKWFSYSPVLENDVRFKDLCIEAAEDNGYKYVEAKRNLGGEDFSFYQ /tränslation="MPDKAALBCEDIKKSIEKSTTQLIDSWYGESCTAFEKTYLILDR KLNTYTEILQEYSKALYDISETYQNTDGDIAKAMSKAVNEASEQK" /tränslation="mnkrivSmvaglsiifttgfvthisAaankaaasvqsqgykmdn akigaggyvpavifnktekdliyartdmggayrwdkannkwifitdgfsdwtmlgces /product="Protein related to MIFH/DOPD protein family, 'transī_table=11 'product="Predicted membrane protein; CF_1 family"

```
/product="SAM-dependent methyltransferase"
/protein_id="AAK7895.1"
/db_xref="G1:15023830"
/db_xref="G1:15023830"
/ccaref=stalababawwyFelbaswwvIRSEYFEERLKYKVLSKVSIKDKITADF
/ccaref=stalababawwFelbaswawikEbkKSAHDKKLNNIYLLKSSLTSVTLFDE
SIDAVFINMALHHIKDAEKAIKEMYRVLKKDGTVVISDVSEHNGKWAREEMFDEWLGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     829 GTTACTAATGAGGAAAGACATGAGGAGTGAGTTAAAGGACAGCAAAAGGATATCTT 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 ThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyBheValAspGlylle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 AGCCAAGGATATAAGTGGATAATGCAAAGATAGGTGCTGGTGGATATGTTCCAGCAGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 ATATTTAATAAACAGAAAAGATTTAATATATGCTCGTACAGATATGGGAGGAGCATAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     475 CAGGATAAAGGTAACACATGGAAAAGATACCAACTTCCATTTAAGGTAGGAGGTAATAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        123 ProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     163 GlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyrGlnSerAspileGinGlyValValTrpValAlaPheAspLysSerSerSerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TATACAGCAGATAAAGTTGGAGTAGTTTGGGAAACCTTTGATCCTTCAACAGGAACAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 TrpGlyTyrAsnGlyValValSerIleAlaAlaAspProlleAsnThrAsnLysValTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83 AlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415 ATAGCAGCAGGTTTGTACACAAATGATTGGCAAGATGAAAATGCATATATTCTTAGTTCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103 SerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   143 PheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyGlnAlaSerLysThrIlePheValGlyValAlaAsp---ProAsnAsnProValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      crredadciagaadredraredacrrredaaaacreaadacraredacaaacaredrere
                                                                                                                                                                                               SNEQIDMWLKNVGFKNIDIENTDLKCKGTSSRGEHTETSIFIAKAAK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12732
377
107
240
16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (1-12732)
                                                                                                                                                                                                                                                                                                       10353. .11498
/gene="CAC0930"
/codon_start=1
/trans1_table=11
                                                                                                                                                                                                                     10353. .11498
/gene="CAC0930"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                x AE007608
                                                                                                                                                                                                                                                                             'note="metB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.34e-93
2009.00
65.41%
50.95%
49.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-3 (1-740)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
No.:
                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                            CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                          /product="Predicted membrane protein"
/product="Predicted membrane protein"
/protein_id="AAK78900.1"
/db_xref="G1:15023825"
/tb_xref="G1:15023825"
/translation="MIRFQKRWAIPPLISILIICGDIEKIISFGIEFDDILCNLFFI
GLGIIGGFRYRWYKRQIKKYTELLKEDRIDKIEYIIYDIERVRGVYNTAPFYIDGYKC
KLRFPQYKCKUNDNYEYYLEPIYKKNGEFKDFIAYADKKNDYKKSIESTNEITECEEK
AEKKRRRRRRNBYNDSRMELAKKNGEFKDFIAYADKKNDYKKSIESTNEITECEEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7003. .8043
/gene="CAC0925"
/gene="CAC0925"
/codon garart=1
/transl_table=11
/product="TPR-repeat-containing protein"
/protein id="AAK78901.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db_xref="G1:15021801.1"
/db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /codon start=1
/transl_table=11
/transl_table=11
/product="TPR-repeat-containing protein"
/protein_id="AAK78902.1"
/db_xref="G1:15023827"
/translation="WGKNEBAKEAYKEAIKKYNFLIIKYPYDVKLNFFKANCLRDIED
YEKABELYSTIDLDNDALEVYLMRARNRISLNKYEAAKEDLNTVIKLNPAYRKVIEL
DEGLAEFLSGKSESCYR"
9379. .8675
VKNPSYWCSSGKTLYDRTSATVTVLANGHSYTSLGKDVEKLPALKKAEASSANKTLEA
ANNBEKITHDAMAFRIKIUGINNAKKSVDLUKCKKONGKTNPEPYKNQPYKDKULL
KKDGNKLPQLQQDLLIQQDQLETPEKNTLANGNKYRKSIRDLNKNYRDIKASLYNIPA
GGTLYNTEDVQGSFHPNYFPGEIKPIPEBENTLANGKERR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RSIYLGEMSEGVYYNLALCMESIGELDDAIKWLARALIINDQNIITLNKKAAVLIKIK
KYBALLECYNRVLKIRUDNEBYYHFKAMLLGEMGKIERAFETISGERALLGEQWTLCY
DKAVLFEKQYGEPEKALLCTEKALLFDBGAKEIRAFTIKKGDLIYLKKIDBAKSTYDKILE
FHENNMEGVFAKANLCMLGEHEDAEEFKEIIERLSGEESVIINSYYYRALNIKKWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Product="Hypothetical protein, CF-2 family"
(protein id="NAK78903.1"
db_xref=11:15023828"
ftranslation="MASDETRIDERACDDVIXDLXAVGEKVKETYNLIHLSNRHIEDG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation="MVDRCDIEIDASIFEETIKVYSDSKVRLHDILCDLENELRKMED
FWEGDAKKEFDSTFPGFYSAMKKOCDMLEELIKELKLVKTSFESLDTEMKNLDKK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KGSAVDSITTEYNKIVKMVKEMQDNMGNDVAAVNNIVSSFDAEDKMIKTKFYR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9712. .10335
/gene="CAC0929"
/codon_start=1
                                                                                                                                                                                                                                                       table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7998. .8357
/gene="CAC0926"
7998. .8357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  codon start=1
transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    379. .8675
'gene="CAC0927"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1713. .9012
'gene="CAC0928"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      712. .10335
'gene="CAC0929"
                                                                                                                6345. .6923
/gene="CAC0924"
                                                                                                                                                                   345. .6923
'gene="CAC0924"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7003. .8043
/gene="CAC0925"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998. .8357
gene="CAC0926"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8379. .8675
/gene="CAC0927"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KTRKOKRLIRKO"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
```

셤

ò

셤 δ 셤 à 셤 ઠે

ò

```
Location/Qualifiers
1. .2950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zverlov, V.V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GERMANY
                                                                                                             638
                                                                                                                                                                              658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                         용
                                                                                                                                                                                                                                                                                                                                                                                                        용
         원
                                                8
                                                                     ద
                                                                                                            ઠે
                                                                                                                                         g
                                                                                                                                                                            ò
                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1776
                                                                                                                                              1296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1777 GTÀTGGTCTCCA---GAGGAAGGTGCTAATGCAGCATÀTAGCACAGATAATAAA 1833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1894 AAGAAATTCTATGGTTTCTTAAATGGTAAATTCTATATTAGTACTGATGCAGGTGCAACA 1953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1834 regacacceretreregretrecacagaacaaagerragareagacegretaaareea 1893
                                                                                              TGCGGACCATATGACGGCTCAGATGGACAAGTTTGGAAGTATAATACCAAAACAGAGAA 1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                519
                                                                                                                                                                                                                                                            341
                                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                            380
                                                                                                                                                                                                                                                                                                                                                                                                                                                          400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   480
                                                               GlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThr 281
                                                                                                                            TrpThrArgileSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSer 301
                   TrpProAspThr11e1lePheArgSerThrAspGlyGlyAlaIhrTrpThrArg1leTrp
                                                                                                                                                                                                                                                                                                                                                                                                                        rescriteatresestrasaccesestracaccescasacrestrasacrasacres
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       501 AsnAspArgHisValAlaPheSerThrAspGlyGlyGlyLysAsnTrpPheGln---GlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThr
ProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThr
                                                                                                                                                                                             GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp
                                                                                                                                                                                                                                                                                          rescessargaagaaararacagaastacasargecassraagacerssaaacerararss
                                                                                                                                                                                                                                                                                                                           AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro
                                                                                                                                                                                                                                                                                                                                                                                         TrpLeuThrPheGly---ValGlnProAsnProProValProSerProLysLeuGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     421 AlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   461 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1657 CAAGATAAAGATTGTGGAATTTCCTACGATGGTGGTAAGAACTGGTTCTCAGCTGGAAGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            520 GluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           117 AATATATCAGGAGTTTATAAAGCAGGAACTGTAGCAGCTGGTGGTGGTGCTAAAACTATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                          MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlyalaProLeulleSeralaLeuGlyaspLeuGlyGlyPheThrHisAlaAspValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer
                                                                                                946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            540
                                                                                                                                                             1006
                                                                                                                                                                                                                                                                                            1117
                                                                                                                                                                                                                                                                                                                                                        1177
                                                                                                                                                                                                                                                                                                                                                                                                                          1237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            260
                                                               262
                                                                                                                                                                                                                                                                                                                           342
                                                                                                                                                                                                                                                                                                                                                                                            362
                                                                                                                                                                                                                                                                                                                                                                                                                                                          381
 242
                                                                                                                              282
                                                                                                                                                                                             302
                                                                                                                                                                                                                                                            322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           580
```

셤

g

8

g

ઠ

g

ઠે

ద

ઠે

g

ð

a

à

ò g ò a à g ò 셤 ò 셤 ò

요

à

```
/codon_start=1
/evidence=experimental
/tradel_table=1.1
/product="beta=1.14-xyloglucan hydrolase"
/protein_id="CAE51306.1"
/db_xrsf="G1:37651953"
/db_xrsf="G1:37651953"
/tradelation="MYKKPTSKIRAAVFAAVVAATAIFGPAISSQAVTSVPYKWDNVV
/translation="MYKKPTEXOLIYARADIGGAYRWDPSTETWIPLLDHFQMDEYSYYGVE
SIATDPVDPNRVYIVAGMYTNDWLPNMGAILRSTDRGETWEKTILPFKMGGNMPGRSM
                                                                                                                                                                                                                                                                                                                   2187
                                                                                                                                                                                                                                                                                                                                                                                                            2247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BCT 11-0CT-2003
                                                                                                                                                                                                                                                                       677
                                                                                                                                                                                657
                                                                                                                                                                                                                                                                                                                                                                678 GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeulleAsn 697
                                                                                                                                                                                                                                                                                                                                                                                                                                                         698 AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeu 717
600 Phe---GlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AJS85344.1 GI:37651952

AJS85344.1 GI:37651952

beta-1,4-xyloglucan hydrolase; xghA gene.
Clostridium thermocellum
Clostridium thermocellum
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Direct Submission
Submitted (09-OCT-2003) Zverlov V.V., Fachgruppe f. Mikrobielle
Mikrobiologie, TU Muenchen, Am Hochanger 4, Freising, 85350,
                                                                                                                                                                                                                                                                                                   GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIle
                                                                                                                                                                                                                                                                                                                                                                                                            2188 GATGGAACAGCGGGAATATTCCGTTCAGATGATGAAGGTGCAACTTGGGTAAGAATAAAT
                         619 AlaValProGlyLysGluGlyAspLeuTrpLeuAla---AlaSerSerGlyLeuTyrHis
                                                                                                                                                                             SerThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSerSerAlaValAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="xghA"
/note="belongs to GH74 and contains a cellulosomal
dockerin module"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clostridium thermocellum xghA gene for beta-1,4-xyloglucan hydrolase, strain F7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New cellulosome components of C. thermocellum Unpublished
2 (bases 1 to 2950)
Zverlov, V.V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Clostridium thermocellum"
/mol type="genomic DNA"
/strain="F7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref="taxon:1515"
346. .2887
'gene="xghA"
346. .350
/gene="xghA"
346. .280
```

IGVVWVVPDKSSSTPGNPTKT1YVGVADKNBS1YRSTDGGVTWKAVPGQPKGLLPHHG
ALABNGMLY1TYGDTGCPYDGNGKGQWWKPNTRTGBMIDITPIPYSSBDNRFCFAGLA
VDRQNPD1IMVTSMNAWWPDEY1FRSTGGTWKNINEWGMYPERILHYBIDISAAPW
LDWGTBKQLPEILHYSTDISDDFNSOBTWKYTGAT1YGCDNLTDWDRGGKVKI
EVKATGIEECAVLDLVSPPEGAPLVSAVGDLVGFVHDDLKVGPKKMHVPSYSSGTGID
YAELVPNFMALVAKADLYDVKKISFSYDGGRNWFQPPNSVGGGSVAVAADAKSV IWTPENASPATTDNGNSWKVCTNLGMGAVVASDRVNCKKFYRFYNGKFYISTDGGLT FTDTKAPQLEKSVNKIKAVPGREGHWLAAREGGLWRSTDGGYTFEKLSNVDTAHVVG FGKAAPGQDYMAIYITGKIDNULGFRSDDAGKTWVRINDEHGYGAVDTAITGDPRV YGRYJATNGRGIYVGEPASDREPVPTPPQVDKGLVGDLINGDNRINSTDLTLMKRYILK SIEDLPVEDDLWAADINGDGKINSTDYTYLKKYLLQAIPELPKK"

140 992 TTTGATTATACCAAAGACATTATTGGAGTAGTCTGGGTTGTTTTTGACAAGAGCAGCAGT 1051 220 577 697 100 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120 817 877 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160 937 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180 241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260 637 757 40 9 80 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAspAsnIle ThrThrGlnProTyrThrTpSerAsnValAlalle---GlyGlyGlyGlyGlyPheValAsp 61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys GATGAGTACAGTTATTACGGAGTGGAAAGTATTGCAACCGACCCTGTGGAATCCGAACCGT |||::: GTTTACATAGTTGCAGGTATGTATACCAACGATTGGCTTCCTAATATGGGAGCAATTCTT CTTTATCTTGGAACACGATGCGGAAACGGACTTTGGAGAAGTACCGACTACGGTGTAACA ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 41 MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp SerLeuGlyGlnAlaSerLysThrllePheValGlyValAlaAspProAsnAsnProVal GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly 2950 387 100 229 24 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps: US-09-917-376-3 (1-740) x AJ585344 (1-2950) 6.26e-93 1982.00 65.81% 52.30% Percent Similarity: Best Local Similarity: Alignment Scores: Pred. No.: 518 21 81 698 101 758 121 818 141 878 161 181 Query Match: DB: ORIGIN

셤

Š

ઠ 셤 셤 ð 셤

ò

원

ઠ

심

ò

심 ò

ઠ

엄

8

셤

ð 셤 8

ઠ

d

qq		
ò		261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSer 279
qq		1229 ACCTGCGGTCCTTATGACGGAAAAGGTCAGGTTTGGAAGTTCAATACACGTACA 1288
\$ a		280 GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly 299
ò		00 TyrserGlyLeuThrileAspArgGlnHisProAsnThrileMetValAlaThrGlnIle 319
q		1346 TTTGCAGGACTTGCAGGAGCAGGATCCTGACATTATAATGGTAACTTCCATGAAC 1405
හි ට		20 SerTrpTrpProAspThr1lellePheArgSerThrAspGlyGlyAlaThrTrpThrArg 339
g (06 GCGTCGTGGCCGGATGAATATTTTCCGCAGTACTGACGGCGAAGCTACATGGAAGAAT 14
음 원		340 ILETTPABDITENTENETTYPTZABRAFGSETLENARGIYYVALLENARGIYVATALLENARGITENETTA 359 1466 ATCHGGGAATGGGGAATGTATCCTGAACGTATACCTGCATTATGAAATAGATATTTCCGCA 1525
ò		360 GluProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGly 379
q		1526 GCACCGTGGCTGGCGAATTGGGGAACTGAGAAACAGCTGCCGGAAATCCGAAACTGGGA 1585
ठे ह	•	380 TrpMetAspGluAlaMetAlaileAspProPheAsnSerAspArgMetLeuTyrGlyThr 399
8 8		GlvalaThrLeuTyrAlaThrAsnAsDLeuThrLvsTrpAspSerGlvGlnIleHis 419
7 음		46 GGTGCAACTATCTATGGTTGTGACAATCTTACTGACTGGGACAGAGGGGGGGAAAGTAAAA 17
ò		20 IleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProPro 439
q		
ò		440 SerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspVal 459
q		1766 GAGGGTGCACCGCTTGTAAGTGCAGTTGGCGACCTTGTCGGTTTTGTTCATGATGACCTG 1825
ò		460 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAsp 479
q		SCACGTT
ò		480 TyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspPro 496
g		1883 TATGCGGAGCTTGTTCCGAACTTTATGGCATTGGTTGCAAAGCCTGATTTGTATGATGTA 1942
ò		7 SerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe 51
Q		
ò		17 GlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAsp 53
셤		s cagcca
Š		536 GlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGly 555
q		2045 GCAAAATCAGTTATTTGGACACCGGAAAATGCAAGTCCTGCAGTTACAACGGAC 2098
à		556 PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAsp 575
옵		2099 AACGGAAACTCATGGAAAGTTTGTACAAATCTTGGTATGGGTGCGGGGGGGG
ò		95
a		racscattctataacsscaaattctatataascacssac 22
ò		596 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGly 614

유 8 셤 ò 원 ò g ò g ò 셤 Š 셤

```
RIGTGDSTSGKQVAISSDYGVTWNQHFGAPDNVQGGKVAISADADIILWRTNGNGVWV
SRNQATENIFYGASGKTFYVSTDNGKTFSAHGSLGSATSVFDITVHPSVSGDIWASTD
KGLFHSTDSGATFSAISGITQAMGVALGAPRSTGGYPAVFAANVGGEVAVFRSDDRG
VNWVKINDAHGFGAASANCWAADREVYGRVYIGTNGRGIFYGDVAGSAPPPTSSATS
TTSVSTTTTRSTTTTRSGNGSPVSGFGQCGGGGWTGGTTCVTGFQCVAQNP
FYSQCIPI"
                GNLQ1GWMGSLS1DPFDSNHWLYGTGET1YGSRDLLKWDSAHNVT1KSLADGVEETS
VQAL1SPPSGPPLVSA1GDVGGFVHQSLTRAPSAQFTNPKWSTSAD1DFAGNNPSNVV
 vaalnswwpdgoifrstnggaswtplwdwavyptlnkyysynaalapwigpnivdvtp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              511 ATCCTTTATTTTGGCGCTCGCAGTGGTAACGGTCTGGAAGTCAACAAATTCGGGTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                     20 AspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 TrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuargSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    crcararcaagrearraregraaaacarrceareceececcrrrecearreaaecreeer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyasnMet ProGlyargGlyMet GlyGluargLeuAlaValAspProAsnAsnAspAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ::|||||||:::::||||:::|||:::||||:::||||
571 AGCTGGACCAAGGTAACAAGCTTCACTGAACACTGGGTCTTTCGTTCCCGATCCATCAGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              180 ThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   200 SerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---AsnAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2579
341
112
249
41
                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                   (1-2579)
                                                                                                                                                                                    97. .2445
/gene="cel6"
/product="CBL6"
                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740) x ABI292929
                                                                                                                                                    28. .96
/gene="cel6"
97. .2445
                                                                                                                                                                                                                                                                                         2.85e-81
1753.50
60.97%
45.90%
43.45%
                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                       sig_peptide
                                                                                                                                                                                         mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                      н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       219
                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                 No.:
                                                                                                                                                                                                                                           DRIGIN
                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      충
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                             2509
                                                                                                                                                                                                                                                                                                                                                 2569
                                                                                                                                                                                                                                                                                                                                                                                                                    2629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /protein_id="CAC02964.1"
| db_xref="ed1:975897"
| db_xref="ed2:975897"
| db_xref="GOA.094T8"
| db_xref="SPTREMBL:09P4T8"
| db_xref="SPTREMBL:09P4T8"
| db_xref="SPTREMBL:09P4T8"
| drainslation="WPSRYKEWITLIGTFSLIVYGQAASQSYRWQNVXIGGGGGFVPG
| IVERPSEKGIAYARTDIGGAXKIANDDFWTFBLLDFADDSRRYWGVDALATDFVQPNR
| IVERPSEKGIAYARTDIGGAXKIANDDFWTFBLLDFADDSRRYWGVDALATDFVQPNR
| IVERPSEKGIAYARTDGGAXKIPDAAPLPFKLGGNMPGRGMGERLVVDPNLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SILYFGARSGNGLWKSTNSGRSWSKYTSFTDTGSFVPDPSDSTGLNSDKIGIAWYTD
KASGSSGSATPRIFVGVANKGSNSIYISTNGGSSWSAVAGQPTSFLPHKGVLSPSERA
LYITYSDGGGPYDGTSGAVYKYSIDTGAWTDITPVSGSDLFFGFGGLALDTQKNGTVM
2272
                                                                                                                     2390 GCTCATGTGGGTAGGCTTCGGAAAGGCAGCACCGGGACAGGATTACATGGCGATTTACATT 2449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-AUG-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Agaricus bisporus
Agaricus bisporus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
                                                                                                                                                                       673
                                                                                                                                                                                                                                                                                                                 713
                                                                                                                                                                                                                                           693
                                                                                                                                                                                                                                                                                                                                                                                   733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                bisporus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ω
                                                                                                                                                                                                                                                                                                                                  2510 GTGCGTATCAACGACGACGACGGATATGGCGCTGTTGATACTGCAATAACAGGTGAC
                                                                                                                                                                                                                                                                                                                                                                                 2450 ACCGGTAAAATTGACAATGTTTTAGGATTCTTCCGTTCCGATGATGCCGGCAAGACATGG
                                                                                                                                                                                                                                                                                                                 694 ValLeulleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAsp
 GGTGGATTAACCTTTACCGATACAAAGGCACCGCAGCTTCCCAAGTCGGTTAAC----
                                 ValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAla---SerSer
                                                                  ---aagataaaagccgraccggaagaaagaaagacatgtarggcrigcaagaaggc
                                                                                                     GlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSer
                                                                                                                                                                         AlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVal
                                                                                                                                                                                                                                           674 ValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      King,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Direct Submission
Submitted (07-AUG-2000) Morales-Almora P., Microbiology, King
College London, 150 Stamford Street, London, SEI 8WA, UNITED
KINGDOM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Thurston, C.F.
the cellulolytic genes in Agaricus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA
protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /function="putative cellulase"
/codon_start=1
product="CEL6 protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. 2579
/organism="Agaricus bisporus"
/mol_type="mRNA"
/strain="D649"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:5341"
/clone="PMRT152"
/dev stage="mycelium"
/country="United Kingdom"
1. .2579
/gene="cel6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Agaricales; Agaricaceae; Agaricus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2579 bp
for CEL6 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Agaricus bisporus mRNA
AJ292929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AJ292929.1 GI:9755896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Morales-Almora, P. and
Molecular analysis of
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28. .2448
/gene="cel6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 (bases 1 to 2579)
Morales-Almora, P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            celé gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABI 292929
                                   615
                                                                                                                                                                         654
                                                                     2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
ABI292929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCUS
```

14

330

79

153

213

59

390

66

139

450

```
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AB015511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                        용
            셤
                                        ò
                                                                셤
                                                                                               ઠે
                                                                                                                         셤
                                                                                                                                                     ઠ
                                                                                                                                                                            g
                                                                                                                                                                                                           충
                                                                                                                                                                                                                                                                 ઠ
                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1569
                                                                                                                                                                                 ||||::: ||||||| :::||| :::||| GGTTTCGGAGGTCTGTTGCTGCCCTC 1041
                                                                                                                                                                                                                                                   1042 AACTCGTGGTGGCCTGATGGTCAGATATTCCGCTCGACCAACGGGGGTGCATCTTGGACT 1101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1735 caggccaccrrcaararrircracgagccrccgcaaaacarrcracgrrrccacadad 1794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThr 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  575
                                                                      318
                                                                                                                                                                                                                                                                              339 ArgileTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSer 358
                                                                                                                                                                                                                                                                                                                                                     CTGGCGCCGTGG-----ATAGGACCCAACATGTTGATGTCACCCCGGAAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThr
                                                                                                                           GlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGln
                                                                                                                                                                                                                                                                                                        ccecrerideacreridecagneriacceacrereacaacaaracrideagranaacecreec
                                                                                                                                                                                                                                                                                                                                                                                           ---LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyPhelleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThr
                           AGCTTCCTGCCACATAAGGGCGTTCTATCCCCTTCTGAGAGGGCACTTTATATCACTTAC
                                                                                                                                                                                                                                                                                                                                                                                                          1210 CTACAAATCGGGTGGATGATGGAATCACTATCCATTGATCCTTTCGACTCGAATCATTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1621 AATCAGCAFTTTGGAGCTCCTGATAATGTTCAAGGCGGAAAGGTTGCCATCTCCGCTGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CACAACGTCACCATCAAGTCGCTCGCAGATGGTGTCGAGGAAACCTCCGTCCAAGCCCTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      516 PheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCCGACATCATCCTCTGGCGTACAATGGTAACGGT---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ArgvalAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAsp
                                                                                                          SerGlyThrTrpThrArglleSerProValProSerThrAspThrAlaAsnAspTyrPhe
                                                                                                                                                                                                                        319 IleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThr
                                                                                                                                                                                                                                                                                                                                     AlaGluProTrpLeuThrPheGlyValGlnProAsnProProValProSerPro----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGly
                                                                                                                                                                                                                                         ::
                                                                                                                                                                                                                                                                                                                                                                                           377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             436
                           811
                                                                                871
                                                                                                                                                                                              982
                                                                                                                                                                                                                                                                                                          1102
                                                                                                                                                                                                                                                                                                                                     359
                                                                                                                                                                                                                                                                                                                                                               1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          576
                                                                                                            279
                                                                                                                                                                  299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
 8.6
                                                     Š
                                                                            셤
                                                                                                          ઠ
                                                                                                                                    g
                                                                                                                                                                  ઠ
                                                                                                                                                                                           g
                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
Submitted (12-JUN-1998) Motoo Arai, Osaka Prefecture University, College of Agriculture; Gakuencho 1-1, Sakai, Osaka 599-8531, Japan (E-mail:motoo@biochem.osakafu-u.ac.jp, Tel:81-722-54-9465, Fax:81-722-54-9465)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        мвильэл
Aspergillus aculeatus mRNA for Avicelase III, complete cds.
AB015511
                                                                                                                                                                                                                                          FNPSAKGVAYARTDIGGAYRLNSDDTWTPLMDWVGNDTWHDWGIDALATDPVDTDRVY
VAVGMYTNBOPNVGSILLRSTDGGDTWTETKLPFRVGGNMGRGKGMGERLAVDPNKNSI
LYFGARSGHGLMKSTDVGATWSNTYFTYFQDSSSTYTSDPVGTAWPTPDSTSG
SSGSATPRIFYGVADAGKSVFKSBDAGATWAWVSGEPQYGFLPHKGVLSPEEKTLYIS
                                                                                                                                                           1969 GGTGTCGCTCTTGGTGCACCTAGATCTACAGGAGGATATCCTGCTGTTTTCGCTGCCGCC 2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2029 AATTACGGCGGGGGGGTAGCCTACTTTAGGTCTGATGACCGGGGAGTCAACTGGGTAAAG 2088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Burotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases I to 2745)
Avai, M., Takada, G., Kawaguchi, T. and Sumitani, J.
Avicelase III from Aspergillus aculeatus
Published Only in DataBase (1998)
2 (bases I to 2745)
Axai, M., Takada, G., Kawaguchi, T. and Sumitani, J.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein_id="BAA29031.1"
/db_xref="GI:3242655"
/translation="MARSSLALLCAALLGKLADAAASQAYTWKNVVTGGGGGFTPGIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YANGAGPYDGTNGTVHKYNI TSGVWTDI SPTSLASTYYGYGGLSVDLQVPGTLMVAAL
                                                            2149 CGTGTATACGGACGCGTTTATATGGGAACCAATGGTCGCGGTATCTTCTATGGTGATGTT 2208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      714
596 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal 615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            656 AsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           676 ThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCygGlyThrThrTrpValLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      696 IleAsnAspAspGlnHisGlnTyrGly---AsnTrpGlyGlnAlalleThrGlyAspHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           715 AlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle
                                                                                                                        616 MetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Aspergillus aculeatus"
/mol_type="mRNA"
/db_xref="taxon:5053"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /product="Avicelase III"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3. .2578
/gene="avill"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /gene="avillI"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AB015511.1 GI:3242654
Avicelase III.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Aspergillus aculeatus
Aspergillus aculeatus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . .2745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        . .2745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2209 GCTĠĠŤTCC 2217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    735 GlyGlyAla 737
```

NCWWPDELIFRSTDSGATWSPIWEWNGYPSINYYSYDISNAPWIQDTTSTDQFPVRV
GWWYDALLIDPPDSNHHIYGTGLITYGGEBLUYNYSGLAVGIEBRANLGLI
TPPGGPALLISAVGDGGFYHSDLDAAPNQAYHTPTYGTTNGIDYAGNKPSNIVRSGAS
DDYPTIALSSNPGSTWADDYAASTGTGAVALSADGDDYLLANSSTSGALVSKSAS
DDYPTIALSSNPGSTATSTGTGAVALSADGDDYLLANSSTSGALLVSKSQGTI
TAVSSLEPSGAVIASDKSDNTVFYGGSGAIYVSKNPATSFTKTVSLGSSTTVNAIRAH
PSIAGDVWASTDKGLWHSTDYGSTFTQIGSGVTAGWSFGFGKASSTGSYVVIYGFFTI
DGAAGLEKSEDAGTNWQYISDASHGFGSGSANVVNGDLGTYGRVFGHERRPGHILRQS
QREPAGRHGDGDGDTTTSKTSTTVSTTLKTTSSASTTSKSSTTVKTTTSSSSTTSKAS

TITKTTTTSSGTTATSASAYAQCGGNGWTGATVCFTGYTCTYSNAFYSQCVPS

ద

ઠે 셤 ठे ద

ℰ

ò

성 암

g

5 8

ద

8

8 8 8

셤 ઠે 요양 g ઠે g 8 g 8 요 ò g 셤

ò

829

&

236 355 116 415 136 475 156 535 176 595 196 649 216 709 769 255 187 295 244 13 33 59 16 96 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIle---GlyGlyGlyGlyPheVal CCCGGTATCGTCTTCAATCCCTCCGCCAAAGGCGTGGCGTACGCGCGCACGGATATCGGA GlyMetTyrArgTrpAspAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGly AATGATACCTGGCATGACTGG------GGCATCGATGCCCTTGCCACGGATCCGGTC LysleuGlyGlydsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsn AsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAsp TATGGACCTGGTCGAATGTCACCTCCTTCACGTGGACTGGAACGTATTTCCAGGAT LysSerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro -----TrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIle AsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAsp GlyAla11eLeuArgSerSerAspGlnGlyAlaThrTrpGlnI1eThrProLeuProPhe SerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsn AsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAla ProThr -- GlyPhelleProHisLysGlyValPheAspProValAsnHisValLeuTyr AspGly1leValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGly AAGGTTGGTGGGAATATGCCCCGGTCGCGAATGGGTGAGAGATTGGCGGTCGACCCGAAC ProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAsp Gecaagrogicircaagrogaagaracagaracaaracraacrogacrogaracrogaaga 2745 353 114 250 33 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps: (1-2745)x AB015511 1722.00 62.27\$ 47.07\$ Percent Similarity: Best Local Similarity: Query Match: US-09-917-376-3 (1-740) Alignment Scores: 20 128 40 188 9 245 77 967 356 117 416 137 476 536 177 596 197 217 710 237 97 157 No.: DRIGIN 셤 엄 셤 ò ð 임 ò g ð 셤 ò 엽 ò ઠે 엄 ò 셤 8 ર્જ 음 ઠે \$ A à

 256	IlealaThrSerasnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPhe 275
 276	SerValThrSerGlyThrTrpThrArglleSerProValProSerThrAspThrAlaAsn 295
 296	AspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetVal 315
 316	AlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyGlyAla 335
 336	ThrTrpThrarglleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeu 355
 356	ASDIJESERALAGIUPrOTrpLeuThrPheGlyValGlnProAsnProProValPro 374
 375 1178	SerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg 394 :::::
 395	MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer 414
 415	GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp 434
 435	ProProSerGlyAlaProLeuIl ccGcccGccGcAccGGcGcTGCT
 45	LePheThrSerProValPheThrThr 474 :::
 475	GlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPhe 494
495	ProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsn 514
 515	nGlySerGluProGlyGlyVal AGACTACGCCGCGTCAACCAGC
 535	AspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaVal 554
 555	GlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSer 574
 575	_ ₹ წ
 595	AspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGly 614

```
3.27e-74
1620.00
55.57%
42.62%
40.14%
                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1691
                                                                                                                                                                                                                                                                                                                                                                                                                                 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64
                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                            No.:
                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ď
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                      AYU4U839 3959 bp DNA linear PLN 01-APR-2002
ASpergillus niger endoglucanase C (eglC) gene, complete cds.
AY040839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /translation="MRRTFAALVAGYLLDSVHAAASQAYTWKNVVTGGGGGFTPGIVF
NPSAKGVAYARTDIGGAYRLNSDDTWTPLMDWANNSNWHDWGIDAIATDPVDTDRVYV
                                                                                     1981
                                                                                                                                                                                                    2101
                                                                                                                                                                                                                                               2102 CAGGTGATCTCGGATGCGTCGCACGGATTCGGCTCG---GGGTCGCCAACGTGAAC 2158
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Aspergillus niger
Bukaryota; Pungi, Ascomycota; Pezizomycotina; Burotiomycetes;
Burotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
I (bases 1 to 3959)
                                                                                                                 673
                                                                                                                                                                                                                                                                                       731
                                                         653
 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Micro-organisms, Wageningen University, Dreijenlaan 2, Wageningen 6703HA, The Netherlands
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hasper, A.A., Dekkers, E., van Mil, M., van de Vondervoort, P.J. and
                                                                        Gecresagerrriegenandecarcercenceagecachacercerantacege
                                                                                                                                                                                                                               694 ValLeuileAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAla-----IleThr
                             GCCATC---CGCGCGCATCCCTCCATCGCCGCGATGTCTGGGCGTCGACGACAAGGGC
                                                         LeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSer
                                                                                                                654 AlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVal
615 ValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGly
                                                                                                                                                                        674 ValGlyThrileGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrTrp
                                                                                                                                                                                                                                                                                     GlyAspHisAlaAsnLeuArgArgVal-TyrIleGlyThrAsnGlyArgGlyIleValTy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .1858,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          major
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .3959
/organism="Aspergillus niger"
/mol_type="genomic DNA"
/mol_type="genomic DNA"
/db xref="taxon:5061"
<1038. .>3841
/gene="eglC"
join(<1038. .1137,1185. .1329,1373. .1571,1618.
/gene="eglC"
/gene="eglC"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of Industrial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .1571,1618.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EglC, a new endoglucanase from Aspergillus niger with activity towards xyloglucan Appl. Environ. Microbiol. 68 (4), 1556-1560 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="endoglucanase C"
join(1038. .1137,1185. .1329,1373.
1908. .1961,2007. .3841)
/gene="eglC"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Direct Submission
Submitted (18-JUN-2001) Molecular Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 (bases 1 to 3959)
Hasper, A.A., Dekkers, E. and de Graaff, L.H.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                         2219 CGCCAATCCCAGCGGAGCCTGCCGGGC 2246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="endoglucanase C"
                                                                                                                                                                                                                                                                                                                                               rGlyAspileGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /protein_id="AAK77227.1"
/db_xref="GI:19879406"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AY040839.1 GI:19879405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Aspergillus niger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Graaff, L.H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appl. Env
21914059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11916668
                             1865
                                                         635
                                                                                                                                             1982
                                                                                                                                                                                                                                                                                                                                               731
                                                                                                                                                                                                                                                                                       712
                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
AY040839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
                             셤
                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                               Š
                                                                                  셤
                                                                                                                 ઠે
                                                                                                                                            셤
                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                       8
 ઠે
                                                         ò
                                                                                                                                                                        ò
```

SGSPTPRIFYGVVDTGESVFVSEDAGETWTWVSGEPWYGFLPHKGILSPSEHTLYISY
SMAGAPYDGATWGTVHKXVIITSGWWTDLSPSSWTATAVYSOGGELAVDLQVPGTVWYAALN
CWWDDELIWRSTDSGGTWSPIWAMNGYPSINYYYSYDISNAPWLQDDTSTDEFPVRVG
WWYDELIWRSTDSGGTWSPIWAMNGYPSINYYYSYDISNAPWLQDDTSTDEFPVRVG
WWYEALAIDPPDSDHWLYGTGETIYGGHDLQNWDSEHNVTIESLAVGIEEMAVLGLIT PPGGPALLSAVGDDGGFYHTSLTTAPSQYYÄTPTYSSTNGIDYAGNKPANIVRSGSSD SPGTLALLSSSFGESWYADYAAGSSTAFGQVALGADADTILLIANSDGAFKSANSATULSA VSSLPSGAVIASDKANITYFYGASGSSFYLSSDTAATFTVTTLGSSTTANAIRAQPS LAGDVWVSTDTGLFHSTNYGKSFTQIGSGCTEGWSFGFGKPSSDGDYPVLFGFFTVDG VTGLFKTEDQGVNWQIISDAEHGFĞSASANVVNGDLQNYGRVFVGTNGRGIFYGDPSG TLPSATATASSASSTAVKSSTGSTSTSKVGSSTTVSSSTATTITTSSIKSTTLTTTKS SSSTTSTSSTATGTASAXGQCGGSGFTGPTQCPSGWTCTYENEYYSQCKSIPGIATDR G"

1330 1390 1510 1570 1690 1450 1273 1154 AATATGAGGGATTGACAAGCTAATTGATCAGGCGGCGGGGGCTTCACGCCGGGTATCGTC 1213 109 129 134 154 174 194 43 63 69 89 23 13 -----GlyValVa 1331 TACGTAAGCTTCTCTTACGAGATGCCTAGGCTCACTGTCCAGGCACGATTGAGGCATTGA pglnileThrProLeuProPhelysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGl 1571 GGTAAGTCCTATATCTGGGTTTTGCTAGCATACACTGACCGCTGCAGCGTCTGGCGCTGTGG SpProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgS erThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrI 174 leAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValA 44 TrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrp 69 lSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrTh rAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTr ----ArgleuAlaValA PheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspileGlyGlyMetTyrArg -----GlyGlyGlyPheValAspGlyIleVal 1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIle-----352 107 255 113 Length: Matches: Conservative: Mismatches: ------GlyTyr-Asn------Indels: Gaps: US-09-917-376-3 (1-740) x AY040839 (1-3959)

648

180 588 219

708 239

FEATURES

```
1180 ATCAAGCGCCTCGGCTGGATGATTGAGTCTCTCGAGATTGACCCAACCGACAGCAGCAACCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           340 ileTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerTrpTrpProAspThr1le1lePheArgSerThrAspGlyGlyAlaThrTrpThrArg 339
                                                                                                                                                                                                                                                                                                                                                arctactttggtgctcgctcagaaacggctctggaagtctacgacggcggcgtgacc
                                                    TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr
                                                                         :::|||::::::|||:::::|||
TTTTCCAAGGTCTCGTCGTCGTCACGGCAACTGGGACGTACATCCCAGACCCGAGTGATTCC
                                                                                                                      ThrGlyTyrGlnSerAsplleGlnGlyValValTrpValAlaPheAspLysSerSerSer
                                                                                                                                         SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---AsnAsnPro
                                                                                                                                                                                                                        Accacciciosaccacercregráricarionidecacecicariaacarcacrecraca
                                                                                                                                                                                                                                                      ValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly
                                                                                                                                                                                                                                                                                          Gretargradecaccaargecedecreaceredecagasargerediandecedecagecase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGAACTTGGAAAGACATCACCCCTGTCTCTGGATCAGATCTA-----TACTTTGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---ProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  395 MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGlyGlyPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly
                                                                                                                                                                                                                                                                                                                           PhelleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrHisAla-----AspValThrAlaValProSerThrIlePheThrSerProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 360 GluProTrpLeuThrPheGly----ValGlnProAsnProProValProSer----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1060
                                                                                                                                                                                                                                                                                                                                                             769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1480
                                                      191
                                                                                     529
                                                                                                                      181
                                                                                                                                                                                         201
                                                                                                                                                                                                                          649
                                                                                                                                                                                                                                                           220
                                                                                                                                                                                                                                                                                            709
                                                                                                                                                                                                                                                                                                                                                                                              260
                                                                                                                                                                                                                                                                                                                                                                                                                               829
                                                                                                                                                                                                                                                                                                                                                                                                                                                              280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                              셤
                                                                                                                    ò
                                                                                                                                                    요
                                                                                                                                                                                         8
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                           GGATSRIFVGTADNITASVYVSTNAGSTWSAVPGQPGKVFPHKÅKLQPAEKALYLTYS
DGTGPYDGTLGSVWRYDIAGGTWKDITPVSGSDLYFGFGGLGLDLQKPGTLVVASLNS
WWPDAQLFRSTDSGTTWSPIWAWASYPTETYYYSISTPKAPWIKNNFIDVTSESPSDG
                                                                                                                                                                                                                                                                                                                                                                                                            LIKRLGWMIESLEIDPTDSNHWLYGTGWTIFGGHDLTNWDTRHNVSIOSLADGIEEFS
VQDLASAPGGSELLAAVGDDNGFTRARNDLGGTSPOGTWARTSVDYAGNSVK
SVORNGGTNAGTQOYA ISSDGGATHWSIDYAADTSWNGGTVAYSADGDTILWSTASSGVO
RSQFQGSFASVSSLPAGAVIASDKKTNSVFYAGSGSTFYVSKDTGSSFTRGFKLGSAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TIRDIAAHPTTAGTLYVSTDVGIFRSTDSGTTFGQVSTALTNTYQIALGVGSGSNWNL
YRFGTGPSGARLYAGDSGASWTDIQGSGGFGBIDSTKVAGSGSTAGQVYVGTNGRGV
FYAQCTVGGTGGTBSSTKQSSSSTSSASSSTTLRSSVVSTTRASTVTSSRTSSAAGP
TGSGVAGHYAQCGGIGWTGPTQCVAPYVCQKQNDYYYQCV"
                                                                                                                                                                                                                                                                                          /tränslation="MKVSRVLALVLGAVIPAHAAFSWKNVKLGGGGGFVPGIIFHPKT
KGVAYARTDIGGLYRLNADDSWTAVTDGIADNAGWHNWGIDAVALDPQDDQKVYAAVG
                                                                                                                                                                                                                                                                                                                           MYTINSWDPSNGAIIRSSDRGATWSFTNLPFKVGGNMPGRGAGERLAVDPANSNIIYFG
                                                                                                                                                                                                                                                                                                                                             arsgnglwkstdggvtfskvssftatgtyipdpsdsngynsdkoglmwvtfdstsstt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CGCTCGTCAGACCGCGCGCAACGTGGTCCTTCACCAACTTGCCCTTCAAAGTCGGGGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp :::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTGTATGCCGCAGCCATGTATACGAACAGCTGGGATCCGAGTAATGGAGCCATCATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AsnMet ProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAspAsnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AACATGCCAGGACGCGGAGCCGGAGCGTCTGGCTGTCGATCCGGCCAACTCCAACTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TITICATGGAAGAACGICAAGGICGGGGGGGGGGGGTICGICCCGGGAICAICTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal----
                                                                                                                                                                                                          endoglucanase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                      'note="anamorph: Trichoderma
                                                                                                                                                                                                        'note="putative secreted
                                                                                                                                                                                                                                                       protein_id="AAP57752.1"
db_xref="G1:31747160"
                                                                                                      db_xref="taxon:51453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740) x AY281371 (1-2517)
 94304, USA
Location/Qualifiers
                                                                                                                                                                                                                                         product="Cel74a"
                                                                                                                                                      gene="cel74a"
                                                                                                                                                                                         gene="cel74a"
                                                                                                                                                                                                                        codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2.49e-72
1579.00
59.26%
42.99%
39.12%
                                                                                                                                                                       .2517
                                      .2517
                                                                                                                                        .2517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
   ð
     Alto,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
                                                                                                                                         gene
                                                                                                                                                                       CDS
```

ORIGIN

Score:

엄

g ઠે 엄

ઠે

셤 δ 셤 ò

ò

ò

유

8

828

299 939

259

1479

491

511

471

394

434

ò

셤

셤 ð 셤 572

ద ò g ð

```
Adachi, J., Alzawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Fujimura, T., Fukuda, S., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iishi, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Karch, H., Kawagashira, N., Kawai, J., Kawamata, M., Kituchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodam, T., Kojima, Y., Kojima, Y., Kojima, Y., Kojima, Y., Kojima, Y., Kojima, Y., Miura, J., Miyazaki, A., Masuda, H., Marakami, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Mikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Sakai, K., Sakai, C., Sakai, C., Sakai, C., Sakai, C., Sakai, C., Sakai, K., Shinagawa, A., Suzuki, Y., Saito, K., Satch, K., Sagano, S., Sugiyama, A., Suzuki, Y., Suzuki, Y., Tagami, M., Tagami, T., Tomaru, A., Suzuki, K., Tagami, M., Tagami, T., Tomaru, A., Tagawa, A., Takahashi, F., Takahu-Akahira, S., Tanaka, T., Tomaru, A., Yawada, Y., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ramamoco, w.
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
FAIS Genome Sequencing & Ishiblki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kursoaki, T., Kusumegi, T., Lu, M., Marada, J.,
Mizuno, K., Marikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninoi, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Itoh, M., Kagawa, I., Kanagawa, S., Katch, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Shinagawa, A., Shiraki, T.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A., and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rice.
URL.

URL.

URL.

URL.

URL.

URL.

URL.

URLS Rice Full-Length cDNA Project Team: Kikuchi,S., Satch,K.,

Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,

Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,

Chneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and

Yamamoto,M.
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN. Kawai, J., Carninol, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Yoshino, M. and Hayashizaki, Y., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                               Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Direct Submission
Submitted (27-Aug-2002) Shoshi Kikuchi; National Institute of
Submitted (27-Aug-2002) Shoshi Kikuchi; National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-86602, Japan (B-mail:skikuchi@nias.affrc.go.jp,
Tel:s1-29-838-7007, Fax:81-29-838-7007,
This clone is one of the 28K full-length cDNA clones from jäponica
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Science 301 (5631), 376-379 (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (bases 1 to 2724)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  foshimura, A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                      TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AKL10506 2724 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:002-167-F04, full
insert sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,I., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1696 CAGCGCTCGCAGTTCCAGGCAGCTTTCCCTCCGTCTCGAGCCTGCCGGGGCGCCGTC 1755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2091
                                                                                                                                                                                                                                                                                                                                                             1645 TATTCGGCCGACGGCGACACGATCCTCGGTCGACCGCCTCGTCCGGC------GTG 1695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1924 TCGACCGACGTCGGCATATTCCGCTCCACAGACTCGGGCACGACCTTTGGCCAAGTCTCC 1983
                  --GIGGCCATCICGICCGACGGC 1584
                                                                                                                                                                                         1585 GGCGCGACGTGGAGCATCGACTACGCGGCCGACACGTCCATGAACGGCGGCACGGTGGCC 1644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2152 AGCACCAAGGTCGCCGGCAGCGGCAGCACCGGGGCAAGTCTACGTGGGCACCAACGGC 2211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 649 ThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr 668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AKI10506.1 GI:32995715
FLI_CDNA; oligo capping.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1756 ATCGCCTCGGACAAGAAGACCAACAGCGTCTTCTACGCCGGCTCCGGATCGACCTTTTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             612 AlaValGlyValMet-----PheHisAlaValProGlyLysGluGlyAspLeuTrpLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1864 AGCGCAGGGACGATCCGGGATATCGCTGCTCACCCGACCACCGCGGGCACGTTGTATGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    687 AspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                        552 TyralavalGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IleargSerAspargValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                592 ArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         630 AlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             669 ProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla-----TyrArgSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2092 gaadacacceccecercercaacacariccaggereccaaggereccaeggerecae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  707 GlyGlnAlaileThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGly
                                                                                               GlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAla
                                                                                                                                                                                                                                                                                532 AlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2212 CGGGCCTCTTTTACGCTCAGGAACCGTCGGCGGCGGCACGGCGGG 2259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              727 ArgGlylleValTyr-----GlyAspIleGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Oryzeae; Oryza.
             1540 GGCAACACCGCCGGCACGCAACAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ehrhartoideae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AK110506
```

1984

셤

ò 유 ò

ò

g

8

ò

g

ò

DEFINITION RESULT 12 AK110506

ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS

ii i

1293 1692 :::||| ||| :::|||||| CTCTGGCAGTGGGGTAACTATCCCGAGATGATCTACTATTACGGCATCTCGACTCCCAAG 1185 ::: |||||||::::::||| :::
TTCGGCGCCTCTCCGTGGACATGAAGAACCAGGAACTCTGGTCGTCGTCGCAGTCTCAAC 1065 1473 ||||||||:: 1474 GCGGCAGACCTTGGAACTTCACCAAAGCGGGTCTGGGACACACCCCAAGTTCACCACCTCA 1533 1590 418 438 515 593 651 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle. SerTrpTrpProAspThr1le1lePheArgSerThrAspGlyGlyAlaThrTrpThrArg IleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla ThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIle HislleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerPro 1534 GCCAGCGTCGACTACGCCGGCAACTCGCCCGGGAAGGTTGTGCGCGCGGGCAAC---GAT gccccccccccag------arcercecarcreaccarcaaraccaccarcacaras PheGlnGlySerGluProGlyGlyValThr ----ThrGlyGlyThrValAlaAlaSer -----AACGTGTACCCCGGCCGACACAAGAACAAGAACGGCGGTGCCGTGTCCTACTCT 534 AlaAspGlySerArgPhevalTrpAlaProGlyAspProGlyGlnProValValTyrAla 554 ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg SerAspargValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSer 594 ThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaVal GluProTrpLeuThrPheGlyValGlnProAsn---ProProValProSerProLysLeu 1186 GCCCCCTGG-----ATCCGCAAGAACTTCCTCGATACCGACTCCAAGTGGCTG GlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGly ---AlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGly ThrservalAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAsp 496 ProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrp GlyvalMet ---PheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer SerglyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThr---GlyVal ProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGlyGlyPheThrHis----

ठे 셤 ò g ò 名

셤 ð

ઠે

```
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/translations/
/trans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /tränslation="MHCPFCRHPDSRVVDSRTTDDGTSIRRRRQCPDCSRRFTTVETC
SLMVVKRSGVTEPFSRTKVINGVRKACQGRPVTEDALAQLGQRVEEAVRATGSAELTT
HDVGLAILGPLRELDLVAYLRFASVYRAFDSLEDFEAAIAELREGRERPAVDDEDHE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GFLRRDPHRPRAYEVRGSDQSSSVQPTDTACKPAASYVPLVGRIAAGGPILAEESVED
VPPLPRQLVGDGELFVLKVVGDSMIEAAICDGDWYTVRRQPVAENGDIVAAMLDGEAT
VKRPKREDGHVWLLPHNSAYQPIPGDEATILGKVVAVLRRV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /translation="MTTTADSATITAQDRSQGRLEPVHAMNEATSHEGPKRSLPPGRPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GIRADSSGLTDRQRRVIEVIRDSVQRRGYPPSMREIGQAVGLSSTSSVAHQLMALERK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product="putative vitamin B12-dependent ribonucleotide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
/transI_table=11
/product="putative SOS_regulatory_protein_LexA"
/protein_id="BAC70174.1"
/db_xref="G1:29606111"
                                                                                                                                                                                                                                                                                           /organism="Streptomyces avermitilis MA-4680"
/mol type="genomic DNA"
/strain="MA-4680"
/strain="MA-4680"
/note="This strain is also named as strain: ATCC
NCIMB 12804 or NRRL 8165.~synonym: Streptomyces
*4 National Institute of Technology and Evaluation
*5 School of Science, Kitasato University
*6 Institute of Medical Science, University of Tokyo
*7 RIKEN, Genomic Sciences Center
Following url is also available.
http://avermitilis.ls.kitasato-u.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DAGAERQGTDRGSGGTVEVPVPATVAD"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /protein_id="BAC70172.1"
/db_xref="G1:29606109"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       complement (3086. .3649)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       complement (5110. .7143)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement (5110. .7143)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             complement (42. .2936)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complement (42. .2936)
                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="SAV2463"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            note="SAV2461"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'note="SAV2462"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           avermectinius
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'gene="lexA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4168. .4947
/gene="lexA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene="ding"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /gene="dinG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1168. .4947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            reductase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Sakati, Y. and Hattori, M.

Direct Submission

L. Submitted (29-MR-2002) Director-General of Biotechnology Center, Dational Institute of Technology and Evaluation, Biotechnology Center; 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bioGenite:go.jp, URL:http://www.bio.nite.go.jp/, Tel:81-3-3481-8424)

This work was done in collaboration with Haruo Ikeda(*1), Jun Ishikawa(*2), Akiharu Hanamoto(*3), Chigusa Takahashi(*3), Mayumi Shinose(*3), Hiroshi Horikawa(*4), Hidekazu Nakazawa(*4), Tomomi Osonoe(*4), Norihiro Kushida(*4), Hisashi Kakuchi(*4), Tadayoshi (*5), Yoshiyuki Sakaki(*6,*7), Masahira Hattori(*1,*7)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis
Nat. Biotechnol. 21 (5), 526-531 (2003)
                                                                                                             2035 AGCGACACGTACCAGATCGCCGTTGGCCTGGCTCT---GGTTCCAACTGG---AACCTG 2088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2206 CTGGCTGGCAAGGTCGCCAACAAGGTATACGTCGGGACCAACGGCGGGGGTC 2265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BCT 10-MAY-2003
                                                                                                                                                                                                                                                                 2089 TACGCCTTTGGCACGGCGCCAACGCCCCAGGCTGTACGGCAGCGCGGACGGCGCGCTCG 2148
                                                                                                                                                                                                                                                                                                                                                                                                             2149 TCCTGGACTGATCTCCAGGCCGACAAGCAGGCTTCGGCATCATCAAC---GGATGCAGG 2205
                                                                                                                                                                                         691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Final finishing process and all annotation were done by H. Ikeda and J. Ishikawa.
*! Kitasato Institute for Life Sciences, Kitasato University
*2 National Institute of Infectious Diseases
*3 The Kitasato Institute
                                                                                                                                                                                                                                                                                                                                            692 ThrTrpValLeulleAsnAspAspGlnHisGlnTyrGly-----AsnTrpGlyGlnAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 710 IleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIle
                                                                                                                                                                                         672 PhevalvalGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThr
                                         SerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Streptomyces avermitilis genomic DNA, complete genome, section
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y. and Hattori, M. Genone, Sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001) 21477403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---IleGlyGlyAlaProSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Streptomycineae; Streptomycetaceae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Streptomyces avermitilis MA-4680
Streptomyces avermitilis MA-4680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299550 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              and Satoshi Omura(*1,*3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AP005031.1 GI:29606108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (bases 1 to 299550)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AP005031 BA000030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        730 ValTyrGlyAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               metabolites
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AP005031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12692562
                                             652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AP005031/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE
PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
```

```
/codon start=1
/transl_table=11
/product=hypothetical protein"
/protein id="Bag701076.1"
/db_xref="G1:29606113"
/translation="MPPTDASTDAEPATSADPARIAVSARSAADSEDTLELRLPDELV
ALFAAEEAAGGEBAAGGGGOGNAVERGEGTAAGGEEDAVGGAEPAGGEDAVGAAFATA
APSVPARAGATAASADSDGCSDAVDPAAPFAAASTVPSARSVVPSAPSVTFAT
                                  /protein_id="BAC70175.1"
/db_xref="G1:29606112"
/db_xref="G1:29606112"
/dr_inslation="MTKPSLPELLHAAVTAVGGTERPGQVTWARAVBAIDDSSHLLV
QAGTGKSLGYLNALAHGBVIVATIALQRQIVERDLPRTVDALHPLLRRRPEF
AWLKGSRYLCLHRALAHGBVYPQDEREGILPQRGIVSRALGQDLLIRLRDWSDFTERGD
RDDLTPGVSDRAWAQVSVSSRECLGATKCAYGAECFAEMSRERAKLAEVVVTNHALLA
                                                                                                                                                                                                                                                                                    IDAIEGAPVI.PQHEVI.IVDEAHELVSRVTGVATGELTPGQVNRAVRRAAKI.VNEKAAD
QLQTAAEGFERLMELALPGRI.EEIPEDLGYALMALRDAARTVISAIGATRDKSVQDED
AVRKQALASVESVHDVAERITNGSEWDVVWYERHDRFGASLRVAPMSVSGI.LREKI.FA
                                                                                                                                                                                                                                                                                                                                                                                                            DRSVVLTSATLKLGGDFNGVGASLGLAPEGTQGDDLPQWKGVDVGSPFDYPKQGILYV
KHLSRPARDGDRADMLDELTELIQAAGGRTLLFSSWRAAGLAAEERSRIPEFPIL
LQGEBTLGELI KNFAADPETCLFGTLSLWQGVDVPGPSCQLVWMDK IP FPRPDDPLMS
ARQKAVEDAGGNGFMAVAATHAALLMAQGAGRLVRASGDRGVVAVLDQRLATARYGSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PLICRWMNDPAVAAFWKIAGDESVTEQHIRAQIGGBORSVPCIGVLEGTPMSYWEIYR
ADLDSLARHYPARPHDTGIHLLIGGVADRGRGIGSTLLRAVADLVLDRRPSCARVVAE
PDLRNTSSVSAFLGAGFRFSAEVDLPDKRAALMVRDRPLRDLL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /product="hypothetical protein"
/protein_id="BAC70177.1"
/db_xref="G1:29606114"
/tb_maref="G1:29606114"
Translation="MNATPISDGGSRTHERAGHDTQLPTVQGPQAVFASEAQAAPAPE
TQAPIAPATQVVPVPENTTPVVQASESVPPPVGRARESERLAGTTADPLDHPBAAQTAA
QAAAVENLLRCWVRENGLVAPDDGTLRIPLLATGTALLVPVHHWSPTGMHRFGLPYLA
DAPQTAPPADAVTVVALLAREAAGSRRAPRPAHSAVDHCAVDHSADLVGRVADSVRRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AVF I DDRRARPADEPDLFLAAEQALLLGHPLHPTRSREGLSEAEARLYSPELRGSFA
LHWLAVAPPVLATNSAWTERGRPVPAEQLTARLAGDGLPMPDGFAALPLHPWQI REIR
HRPEAAALLDAGLI RDLGPHGPAWHPTSSVRTVHRTGAPAMLKLSLGLRI TNSRRENL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RKELHRGVEVHRLLRSGLGEGWQAAHGFDIVRDPAWLAVDDGHGNAVPGLDVMIRHN
PFAFTDDVSGVAGIVSPRPREDPTBOPRDRRLLEIITRLAGRTGRPRGAVATEWF
RTYLEGVVRPVLWLDSBAGIALBAHGONTLLLLDPEGWPTGGRYRDNGSYFFTSRRA
ELEARLPGTGEHSDPTFYSDEVTDBERAYTGINNVLGLIGAFGSGUHABEQLLLAAFR
RFLADVATGPARLRTPLPTRLLDSFVLRCKANLLTRLHGLDELVGFVDTQSVYVTIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="putative class III aminotransferase"
/protein id="BAC70178.1"
/bc.xref="GI:29666115"
/translation="WAVTKGAVASLKREAEGAHAAHEGILRRQSARESAARTYARALP
/translation="WAVTKGAVASLKREAEGHAHPVVLEAIRKVLDSGAPLHVLDLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YVREHRLAERAQTIGARMLDOLRSLATEFPCVGEVRCRGLMICVELVRAEEDREAVSP
GSGDFASLDTAGAGPRPAAPELAAAVQRECLRRGLIVELGGRHASVVRLLPPLTISDE
QATAVLDRLADAVEAVBAVRARGRAERSRGFQRSERAG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YHGMTAGALEASGGATDGRVARLPYPODYRCPFGIGGERGAELGARWTESLLDDYKSG
VPLPAGMILEPVQGEGGVIPAPDAWLRRMRAITADRSIPLIADEVQTGVGRTGRFWAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APAPAPPSLGDTKALAGRQAVSVTRAADDLLGRVEDWGPTVTPAGTFQLAPVRPERDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EHSGVVPDVMVLSKAIGGSLPLAVVVYRDDLDVWQPGAHAGTFRGNQLAMAAGTATLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPVKDAFTTELFRTLPAGLADRARVQFCGPAGTDAVEAAFKLVRTATGRSGLLAFTGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LKASLPDFWYTTDRNQVRRSLTAIDEAAKKTEVTEVTEVAEVTEEAAKESEGA"
complement (7376. .8509)
'product="putative ATP-dependent helicase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon start=1
/transT_table=11
/product="putative secreted protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        complement (10775. .12229)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           complement (8549. .10609)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="SAV2467"
/codon start=1
/trans1_table=11
/product="putative c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            codon_start=1
transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="SAV2466"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      note="SAV2468"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   note="SAV2465"
```

CDS

```
---cegcaceaceactectcaagtceacceaceececectrese
                                                                                                                                                                                                                                                                                         160673 TACGGCGACGCACCCGGCCCGGCCCAGTCCGACGCAGCGTGCACAAGCTGCGTACC 160614
                                                                                                                                                                                                                                             161381 TACCGCTGGCGCAACGCCGTCATCGGGGGCACCGGCTTCGTCACCGGCGTGCTCTTCCAC 161322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160790 TACCGCACGGCCGAC---GGCACGACCTGGGGGCCGTCCCCGGCCGGCCCTCCGGCACC 160734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /protein_id="BAC70179.1"
/db_xref="G1:29606116"
/translation="WRPIRPIPARRGRSARRRTSPVLGAVGLATALALTATACDSGD
/translation="WRPIRPIPARGRSARRTSPVLGAVGAWKONWNRDDWLREAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            166 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185
                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlalleLeuArgSerSerAspGln 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrSerGlyThrTrpThrArglleSerProVal-------ProSerThrAspThr 293
                                                                                                                                                                                                                                                                         45
                                                                                                                                                                                                                                                                                                                                                                                        85
                                                                                                                                                                                                                                                                                                                                 65
                                                                                                                                                                                                                 TyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                         GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161084 Gececcaccrecaccecaccecaccreaccereaagercescecaaceaagaccec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      186 AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerLysThrllePheValGlyValAlaAspProAsnAsnPro-------Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAla
                                                                                                                                                                                                                                                                                                                                                                                        AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerVal
                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                            Indels:
                                                                                                                                                                                      (1-299550)
                                                                                                                                                                                      US-09-917-376-3 (1-740) x AP005031
                                                                                   4.61e-61
1405.00
54.57%
40.79%
                                                                                                                                            34.81$
                                                                                                                Similarity:
cal Similarity:
                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       258
                                                                                                                                                                                                                   G
                                                                                                                                                                                                                                                                                                                                                                                                                                                86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    160964
                                                                                                                                              Query Match:
                                                                                                                Percent Sim
Best Local
                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

Ŷì

. · · · · ·

source

FEATURES

gene

SGS

JOURNAL

```
gene
                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene
                                                                                                                                                                                                                                                                                                                                                    CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /tränslation="mtrslmsmhspvarviassparrrlrplllsalalalalagge
QDDKakrpanvapiagsaattpataaapaaqtpmpqfvskdgkhalmvdgapflilgaq
VNNSSNYPGIMDKVWPAMQVLGPNTVQVPIAWEQVBPEEGRFDFSFVDTLLNQAREHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NQLGRANGERLAVDPHDGRVILLGSRDAGLWRSDDRGAHWARVEGFPADALAGATARN
HVGSEQAVGIAFVVFDAASGHAGAATPRIYYGVSTAQTSLYVSEDAGRTWSAVAGQPK
GLRPSHWVGHNAGQWYLSYGDRPGPDLMAGGALWNYDATQGRWREISPIPQPATGDGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GWGAVAVDPQHPQVLLASTFRRRTPRDEVFRSGDGGRSWVPLLAAAQFDHSAAPWTAH
ATPHWIGALAIDPFDSNHATFVTGYGIWASRNLQTFSEQQPLQWWFQDRGLEETVPLD
LLSPWAGAHLLSALGDIDGFRHDALDTAQLQYLGPRLTNGESIDGAGQAPQWVVRSGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VRDRRNNEIRALYSQDGGTHWAAFASEPPRGQGAGTIAIAADASQVVWPDQGGVWRT
GRGGRWQQRVGGLPDTAVVVADRVDAQRWYAADRVGGRLYESSDGAASFRDTGQQVGS
PARDERARPQIRPDPWRAGVYYLASFTLGVWRWQNGQIRTLSKSDEARSLGJGKALRA
GAPPALYTAGRVAGVOGIFKSDDEGAHWRRINDDAHRFGKPYSYTGDRIAGRYVFAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WKAATPANAGEFTAVGYFFAKELRASTGVPIGIVNSTWGGSAIEAWDAASLGLNADQ
NQGAIEAIKQRDAAAQAATGKRIARWPKVEDEMPQWREAAFDDSDWDSIPVTKQWESS
GYDGMDGIAWYRTITILSAAEAKAGITLGVGQIDDSDTTYVNGQQVGSTEKQWNLPRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VSLTDNKNOLPTLLYNQMIHPLOPFPVKGVIWYQGETNATDTGAVKYRBQFAAMIRQW
RARGONCTIPPTWYQLANFRAGGOKGELSPWALLBESQGYKTLALDATGOAVIIDIGNP
TDIHPTWKRDVGHRIALAARHVAYGETLUYSAPVFKRASFDGGKAVLGPDLQGSALQV
RGGGAVQGFRIAGADQRFHPATAQIDGDRVIVRSDAVAAPVAVRYGWSENPDDANLIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /trānslation="MIAVIDPPTKRRFSAAHAWLTAVCLTLVLLFWTSEASADEPGTP
DPYQWHSVAIGGGGFVTGVLFHPAERDLAYARTDVGGAYRWDAQAQQWYALTDWLGAD
DWNLMGIDAFAVDPADPNALYLAAGTYWHERAGNAAVLRSFDRGKQFERADLPFKIGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="GI:21107969"
/translation="WTASFVRRGCLLGLVLVSPCAWAVPTLPLLLADGAVLQRDQPMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VWGWSSPNAAIAVSFDGKRATVKADATGQWKVRLPAHAAGGPYVLRVQGDGĞELQVRD
VLVGDVWLAGGQSNNEWPLAQASDGPQAVAAANDAQLRQFKVPKSWSVQPQARLTGGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YOVPAAALKAGVNHIAVRVEDLSGGGGMHGPDEQRFVQTSAGAKRALDGWKFRPAAVR
   đe
Submitted (28-NOV-2001) Departmento de Bioquimica, Universidade
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "note="identified by sequence similarity, putative, ORF ocated using Blastx/Glimmer/Genemark"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'note="identified by sequence similarity; putative; ORF
'ocated using Blastx/Glimmer/Genemark"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             _note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
                                                                                                                                                        axonopodis pv. citri str. 306"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             product="sialic acid-specific 9-0-acetylesterase"
protein_id="AAM36635.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     transI table=11
product="conserved hypothetical protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /protein_id="AAM36634.1"
/db_xref="G1:21107968"
                                                                                                                                                                                                                                                                                                                                             /gene="celA"
/note="synonym: XAC1770"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /protein_id="AAM36636.1"
/db_xref="GI:21107970"
                                                                                                                                                                                        /mol_type="genomic DNA"
/strain="306"
                                                                                                                                                                                                                                                      /db xref="taxon:190486"
/note="pathovar: citri"
                                                                                                                                                        organism="Xanthomonas"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
/trans1_table=11
/product="cellulase"
                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2431. .4347
/gene="XAC1771"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2431. .4347
/gene="XAC1771"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4344. .6101
/gene="XAC1772"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RDALPVSPPRTDTW"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4344. .6101
/gene="XAC1772"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            codon_start=1
transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                           /gene="celA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGRGI FYGDPR"
                                                                                                                                                                                                                                                                                                                       87. .2342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_B
/transl_
```

gene

CDS

gene

CDS

```
PGTWAQVPCADDADE I FHAWS IGRY I DQI AQAGKAEY PLPMYVNAALKGP FNPGQPGQY
ASGGPTDNYLDVWKAAGPHI DLLAPDI YMEBYRMYTTVLERY AR PDNALFVAETGRRP
EYARYLY PTLGHNG I GWSAPGI DYTRY SNY PLGAKHVNEETLAP FALGFQA VGWGRRP
FAKRASEGKLHGTNÆEPGQAVQELKLUNDRWSAT ITYGVPGWFKGEPPGUPED I GAAL
I ABLISPDEFL VTGYHARVTLHPASDATANMI YDRVEGGVYEGNDWKFQRNWNGDQTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /processin id="AAM366371"

/db_xref="G1:21107971"

/db_xref="G1:21107971"

/db_xref="G1:21107971"

/dc_amalate.mgrPd0clewrPDRLRRQFFQRAAAAQDQAGNLQISAPATTPRRR

RDGTASPYFCAAA_PAPFHRSKGYTMQTTIRSAGAMPANRTSLAQCLALSLALLANA

AHAQENRKAADGVTVVPSAKGAAPVRLQVVDAGIIRVSADPDGDFARSPSLMRVPVQG

DTAFQVAREQGDSVQLKTGKVTASISTVDGHVSFADANGKPVLSEVAGGRSFAPLNVVBG

KQYLSVRQRFQSPDDBALYFGGQQGGWNQVGRNVELQQNNIDMAVPYLVSSRNYGL

LWDNNSITRLGDPRGLQPLPKTLTLYDAKGRAGALTARYAINGKHILERRESENVNYQY

LSDLTKYPKKAIITKDKNSRMQVTWEGEIEALTGGEHTFSLYSSEYAKLVVDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /trainiation="Mitrrelickatgaalavigavpaaagatgaskttalppvaaba
JQUKYREPREWEALPVGMSRIGAWWGGALARERQIMEDITYAGGPEDSTSPBALA
ALPOWRALIFPRAEKLADAKLISRPIKQMPVQPIGDLILDPDRADGISDYRRQL
DLDTAVATTTFRSGGAVHRREVFVCAQAQCIVVRLSCDRPGGISLRVGIDSPQTGEVT
ARSYQRFDAVDGPAGIEGRLRPALKVLPVQVSGKLSQVRNSLRIDDAADEVULLISA
ARSYQRFDAVDGPAGIEGRLRPAKLARAKLDPPALLRAHIADHGRLFRRVAIDLGSSEAV
QLPTDERVQRPAEGNDPALAALKKAAKLDPPALLAHAHADHGRLFRRVAIDLGSSEAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KYT ININTEMNYWPSEANALHECVEFLEAMLFDLAQYGTHTARAIYDAPGWVVHINNTD
LWRQAGPIDGAQWSLWPWGGWWLLQQLWDRWDYGRDRAYLSKVYPLFKGAAEFFVATL
WRDPOTGAWNTNPSWSPENGHPGAAVCAGPSWBADGLLRDFTRACITAWSKLLGIDAQL
AQQIAALREQLPPNR IGKAGOLQWDGWDWDWQAPEINHRHYSHLYALHPSSQINLRDT
PELLAAARRSLEIRONATGWGIGGRQWDWNQAPEINHRHYSHLYLLLLISPERTYPNLFD
AHPPPQIDGNFGGTAGITEMLLQSWGGSVFILLPALPKAWPRGSVRGIRVRGGASVDLE
WEGGRLQQARLHSDRGGRYQLSYAGQTLDLELDALPKAWPRGSVRGIRVRGGASVDLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WRQNWNPWNHEFKLDLEFGQRHTVKVEWDLIDPSYIALIHRDPLPAAEAKDLSLWSEA
GQMIDYYFVSADSYDQDAVGYRALIGGKSTMLPKWAYGFWQSRERKKSQDGLVGANAEY
RKRKLSLDNIVLDAG
PTTANYKELDAAGFWFKRNVEVGELDWIGKSYKNSFYDPYSEKAQAIYWRQINEKLYS
PTTANYKELDAAGFWFKRNVEVGELDWIGKSYKNSFYDPYSEKAQAIYWRQINEKLNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LVAPVTRFGATSRQVYLPAGSSWLEFATGKRYEGGQSIEAAAPIERMPLFVRAGSIVP
TGPVQEYVDQVADAPLTVVVYTGADGQFSLYEDDGKGYGYEKGEFSRIPLVWNQAKGE
LSIGKREGSWTGMQAKRTINVRFVDGPRDDAGALAPKTDASIQYDGKPVSVLQRKIAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KGFDAWWDADEPDVHSNLDIAERKARTTPNALGSSTEYFNSYPLPHTHGVYVGDRAA
DDKRVFILSRKGYAGTQRNAVAVWSGDIVSRWDDMRDQISGGVNMALSGLPNWTFDIG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GFAVEKRYEDQDPAHLPEWRELNTRWFQFGAFVPIFRSHGQFPYREIWNIAPEGTPFY
ESMAYYNRLRSALLPYIYSLAGDTYQRDGVIMRGMMMDFPNDPKVRDINDQYLFGPAF
VRLVILWFATWKNNGPQYAPHWVKTDNQRFPRVVTREGKSIGSLSPHAQATLDADRKA
FVAFWQHLKKVDPQHTVIMIQPENEAGTYGSVRDFSPWAQKLFDGQVPDELIKRMNKQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note="identified by sequence similarity; putative; ORF coated using Blastx/Glimmer/Genemark"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /product="conserved hypothetical protein"
/protein_id="AAM36638.1"
/db_xref="GI:21107972"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product="alpha-xylosidase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /transl_table=11
/product="D-xylulokinase"
                                                                                                                                                                                                                                                                                                                                                                                                           'note="synonym: XAC1773"
                                                                                                                                                                                                                                                                          GVNFSSVPQLLKIKLATYK"
5999. .9109
/gene="xyls"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
/trans1_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11570. .13063
'gene="XAC1775"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /gene="XAC1774"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="XAC1774"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gene="XAC1775"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /codon_start=1/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .13063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9201. .11573
                                                                                                                                                                                                                                                                                                                                                                                                                                                 5999. .9109
/gene="xylS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3201. .11573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GKARRR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11570.
```

```
1842
                                                                                                                                                                                                                                                                                                                                                                                         355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       475
                                                                                                                                                                                                                         1137
                                                                                                                                                                                                                                                                                                                                                                                                                                               1233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                             ď
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ⋧
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ፟
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                          8
                                                                                                                                                  8
/protein_id="AAM36639.1"
| Ab xref="dd1:21107973"
| Arabalation="MRLYVGLDVGTQSVKLVAXDPQDRAVVATIAAPWELISRDDGTR
| FROLDGEVERNORGISVSGQQHGFVPVAADGSVTAPVKLWCDTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlyTyrGlnSerAspileGlnGlyValValTrpValAlaPheAspLysSerSerSerSer 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyrilealaThrSerAsnThrGlyGlyBroTyrAspGlySerSerGlyAspValTrpLys 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PheSerValThrSerGlyThrTrpThrArglleSerProValProSerThrAgpThrAla 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPhe 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArg 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsn 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TACTTGAGCTATGGCGATCGCCCCGGCCCGGATCTGATGGCCGGTGGCGCTGTGGAAC
                                                                                                                                                                                                                                                                                                                                                                                                             ThrThrGlnProTyrThrTrpSerAsnValAlaileGlyGlyGlyGlyBleValAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               rarcecrisseacececasececasereseresecereseceariosrasseseceae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             rcerrcearcecearaecaerrceaacececeaccrecerrcaaecreerecaar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Grericiona de de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de deserción de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <u> Aceccaeacceerarcaerescaceseseseseseseseseseseserrrercaceseses a contractores de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore del contractore de la contractore del contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la contractore de la con</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TACCTGGCTGCTGCACCTACATGCACGACGCGCCGCCAAT---GCCGCGGTGTTGCCGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProHisLysGlyValPheAspProValAsnHisValLeu---------
                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                (1-14520)
                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-3 (1-740) x AE011809
                                                                                                                                                             4.22e-50
1162.50
51.07%
35.73%
28.80%
                                                                                                                                                                                                                                          Best Local Similarity:
Query Match:
                                                                                                                                                                                                                   Percent Similarity:
Best Local Similari
                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                      N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      924
                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           장염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ર્જ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ર્જ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
```

1496 1436 1553 ---CGCGATCGTCGCAACAACGAGATTCGCGCGCTGTATTCGCAGGAGGAGCACGAACAAC 1670 2007 cereciesecerestrarciósceasecesécerosécerositarecerióseasaaces 2066 334 454 474 335 AlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrVal 354 LeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProValPro 374 495 AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsn 435 LeulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAgpLeuGlyGlyPhe CTGCTCAGCCCGATGGCCGGCGCGCATCTGCTCACCCCGCCTGGGCGATATCGATGGCTTC 455 ThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThr 1497 cggcar---caccccrrcacacaccccacacrrcacacrrcarccrcacccrcaccarr GlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPhe 1554 GCCGAAAGCATCGATGCCCCCGCCCCCCCCCCCCCCGCTGCTCCCCCACCCGTACCGTG 534 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla 554 ValGly---PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIle ArgSerAspArgValAsnProLysThrPheTyrAlaLeuSer-----AsnGlyThrPhe TyrkrgSerThrAspGlyGlyValThrPheGlnProValAla-----AlaGlyLeuPro SerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg 609 SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLySGluGlyAspLeuTrp GCCAGCACGTTCCGCCGCCGCCGCGCGCGATGAGGTGTTTCGCAGCGGCGATGGCGGA 415 GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp GCGCGCGATGAA------CGTĠĊĠCGCĊĊĠCAACTGCGTCCCĠĂĊCCCŤĠĠ 1020 TACGACGCGACGCAAGGCCGCTGGCGCGAGATCAGCCCGATTCCACAACCAGCCACT---315 ValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyGly LeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSer----295 AsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMet 1197 CGCAGCTGGGTTCCGCTGTTGGCGGCTGCGCAGTTC----------darcacadedececededreg---accedeacecaaacacedeac

Trans.

1.0

us-09-917-376-3.p2n.rge

셤

ઠે

셤

요

ò

셤 ઠે

8

```
RLERUGNTTPERAVIKMEEVMEELYGIKGTLMEGYVUIMKELLKGDFGBEMYYYPASVIS
LIRNSLPWTIGLLTLTTLISWVLGIVIGAILGYYKESFGSKILTGILMAINPIPYYIF
ALTLLLLEVYVYPEIPPLGGGSBVGQIXSLSWNSTLIDVIKMFTBALSLIIIGFANGG
GTKILVQNIREEDYVYARMMGENSKILESSYVLRNAMLPOYTGLGLSLGGIFGGAM
TEILFTYPGIGTLAYNAVRANDYWVMIAISFSIIAVATALLILDLLYPLIDPRIRYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /transT_table=11
/product="01igopeptide ABC transporter, permease protein"
/protein | da" AAD35390.1"
/db_xref="G1:4980802"
/db_xref="G1:4980802"
/db_xref="G1:4980802"
/translation="MKRILER:PEDWMEDLRFKYAVTVLTALVLLSILSIFSPYDPYKW
/translation="MKRILER:PEDWMEDLRFKYAVTVLTALVLASILSIFSPYDPYKW
YVLPTROPPSIENIIGNTGOKGQDIFWLLTFGIKNSLILATIAAVFSRIJATIGMLSG
YLGGTTDRILMALTDTFMVMPVFPLLLLLSSLIKNYLSLPILGIIGVFGWAGDARVI
RSMILSLREREFVTSKFSGMRTFEVVFGDFYPYLIPVISGFIGSMMGAIGPRILA
ILGFTRYBIPTIGSMFHWMINFQAMLLGYWWWVLTPIVTAVFFTALYTLSLSINEYL
DPRTRMQRIGAVK"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /traislation="meklavaenlravyfiksgdkttsiravdnvslhffenevygla
gesgcgkstlikvlegeikpliklvdgevyfkldgesyrinsmslidjikkirrysy
vpgsmsvlnpvrkirrifloviethymsreeafsriesyfealglpvrvldsyphq
lsggmrqrvtiallavtlnekvifalbettaldvvvQrgvijllikkihkeQkvTlvIvV
hdmgvHahlatrigvmyagrlieeaetheIfknplhpytrflinslprfgdrsrkegt
                                                                                                        GVGSGWWKHLPDVAECLLLSYGFKRDKNGKWILPNGTPWKIQLGTNPNDLVAFCVAQQ
WREFGIDVDVITTYISGINEGNFEVLITGNWPAPEPWGGHADLHRTLYFFHSKYQNKP
LGEGVGHVSRWFDERLDKIIDEMEKLDWNDPRNIEFGREGLKILAEELPTIFIGTIGV
LTTALVYNDYYWTNWPTLENPYMMPYWHWSQFKYLLTFLEPTGRK"
                                                 YSDPPFVPLVDPVVGGITINTAKYPPNMKEVRWALALALDIVKLSETEEGNMLISALH
IPPLPLYRKLYYEPLEEWLQNLTVNGVKVYDPDVPLKIAEVAKRYGWNVPDDPEKIKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="similar to GB:U08186 PID:473270 GB:AE000512 percent
identity: 76.00; identified by sequence similarity;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              translation="MKFIKSYLIPRLIQYVYMVVIGITLVFLIPRLSPVDPIEQILNT"
ffdrwgywwpmpkhyfekvedpvgftfnppvgtgpyvlkdydpagywvlwerredw<u>o</u>
                                                                                                                                                                                                                                                                                                                                             /note="similar to GB:Pyro h percent identity: 59.88; identified by sequence similarity; putative" /codon start=1 /transl_table=11 /product="oligopeptide ABC transporter, permease protein" /protein id="ABA35389.1" /db_xref="GI:4980801"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PGSPPSLANL PSGCPFHPRCP1ADQ1CMEMEPDL1EVARNHKVACWKVGVER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="similar to GB:Pyro_h percent identity: 57.09; identified by sequence similarity; putative" (codon start=1 (trans_lable=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="gimilar to GB:Pyro_h percent identity: 65.51;
identified by sequence similarity; putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /producT="oligopeptide ABC transporter, ATP-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /product="oligopeptide ABC transporter, ATP-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /protein_id="AAD35392..1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein_id="AAD35391.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               db_xref="GI:4980803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /gene="TM0302"
3073. 3942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3073. .3942
/gene="TM0302"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4928. .5878
/gene="TM0304"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
/transT_table=1
                                                                                                                                                                                                                                  2050. .3054
/gene≂"TM0301"
                                                                                                                                                                                                                                                                                            2050. :3054
/gene="TM0301"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /gene="TM0303"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /gene="TM0303"
3945. .4931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /gene="TM0304"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3945. .4931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein"
                                                                                                                                                                                                                                        gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                              3187 TTCCGCTCCGACGAGGGGCCCATTGGCGACGATCAACGATGCCCATCGCTTC 2246
                                                                                                                                                                                                                                                                                                                                                                                                                                         BCT 02-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /translation="MKRFLFVAVLFTIFAASFAQFLNVPREQTVIVPGPKTANPGDFN
VIAGWKDPHKGIQQLLLEPLWMDDPVRGEIINALAETGPIYNEDFTKMIVKLRKGVYW
                                                                                     SDGVPFTADDVVFTVEMCMKYPEMARHMERNBYIDKVYKEDDYTVVFELKKPNTRFHT
                                                                                                                                                                                                                                                                 703
                                                                                                                                                                                                                                                                                                                                                                                  704 GlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGly 723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Thermotogae, Thermotogales; Thermotogaceae; Thermotoga. (Dases) 1 to 13278)
Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J.,
Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,
McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.,
Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillipp, C.A.,
Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D.,
Eisen, J.A., Fraser, C.M. et al.
Evidence for lateral gene transfer between Archaea and bacteria
from genome sequence of Thermotoga maritima
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J.,
Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,
McDonald, L., Uterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.
Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A.,
Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D.,
White, O., Salzberg, S.L., Smith, H.O., Venter, J.C. and Fraser, C.M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (01-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="similar to GB:AE000657 percent identity: 48.20;
identified by sequence similarity; putative"
                           -----AlaileThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAla
                                                                                                                                                                                                                                                                 684 TyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyr
                                                                                                                                         664 ProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13278 bp DNA linear BCT 02-J7 maritima section 24 of 136 of the complete genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        periplasmic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="oligopeptide ABC transporter, oligopeptide-binding protein, putative" /protein_id="AAD35388.1" /db_xref="G1:4980800"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="Thermotoga maritima"
/mol_type="genomic DNA"
/db_xref="taxon:2336"
150. 1985
150. 1985
/gene="TM0300"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2304 Accedecedecerescarerrerarescear 2333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       724 ThrAsnGlyArgGlyIleValTyrGlyAsp 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    codon start=1
transI table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AE001712.1 GI:4980799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (bases 1 to 13278)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     maritima
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Thermotoga maritima
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1. .13278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Thermotoga maritin
AE001712 AE000512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Thermotoga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99287316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10360571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AE001712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
```

SOURCE

gene

CDS

```
203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9999
                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         263
                                                                                                                                                                                                                                                 Query Match:
DB:
                            CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                     /product="endoglucanase, putative"
/product="endoglucanase, putative"
/product="endoglucanase, putative"
/protein_id="AAD35393.1"
/db_xref="Goot #AAD35393.1"
/db_xref="Goot #AAD35393.1"
/db_xref="Goot #AAD35393.1"
/translation="MLRSFLILFLAILGVVFGATFEWKSVEINGGGFVPGIIFHPASP
GLLYARTDVGGLYRWDEETKRWKQLPFLRRDGSDYMGYLSVALDPSDPRRIYAMTGK
YTQDMAGYGAILISEDYGETWTIVNLDKYGIKVGGNEDGRRAGERLQVDPNFSSVLFM
GTTKYGLMKSEDFGKNWKKVDSFPSTSVTFYLPDEKSGERGSPTPRIFVGCSEPKGIF
VYEDGGTTWNVLDNLENDLIPLRGKHINGILVYTLSNALGPNGATRGAVWKYVTADOK
NYDVTPWKGDFGYCGIDVQENVVIVSTLDRWYPHDEIFISLSANFDI
NKAPNITALANHNISDVKIDPPDMNRALFTGYGYGWYYTELKKREFEGWRCKYNAIFEN
RGLEETVVLQLVPPTGREFPLLSAILDMGGFRHESLDTPPSSSYKYKPLKWTSLGTARAYO
NSKFVARVHTYTYPPIGSPRPPEDAIDANDMGGFRHESLDTPPSSSYKYKDLKWTSLGTARAYO
NSKFVARVHTYTYPPIGSPRPPPASPAPASPDVNRGFRHESLDTPPSSSYKYRDPGILSKDGKS
FMKGARLEFENWWVSLYSPEDGOINWREIETVPEGITDGGRLSLAVSNDGKTLVWSPA
NHEVLYSDKGKGRTDYPALSAILANDMGGFRHESLDTPBGGKTWMRINNDKHQFFGWIHYMIG
AYVIGFGGAPRKGTDYPALIXANDWNNGYYGIFMSTDEGKTWMRINNDKHQFFGWIHYMIG
DMMEFGRIFLGGTEGGTGTFLYGTES"
/db xref="GI:4980804"
/trānslation="MRVSVRNLTKIFRVGSIFSLKKLVAIDNVSFEIGSGEIFSLAGE
SGCGKTTTAKILLGLEEPTNGEVLYDGIKLKDFVENEMHFLKKVQTVQQNPFSTFNPF
                                                                      RKVEDYLFKTYLRFKWANTKOBALRIJEEKLSVVGLNFDEIKGRYPSEFSGGGLGRIS
IARALLTNESFLVADEPVSKIDASIRIJSIVNLFRDLKERFGTSILYTTHDLSTAYYVA
DKLAIMFRGSIVEMGRAQDILDTPFHPYTKLIKSSIPVPDPPYEWDENINLSSTEREE
FFRAGCKRASRCPEVMSICRESVPPDFDVNGRVVKCMLYK"
5894. .8017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /transl_table=11
/product="L-fucose isomerase, putative"
/protein id="AAD35395.1"
/db_xref="GI:4898090"
/translation="MKGKYRVGLISFSDGREYVHRELEGIVKKFEDEIAKTLEETGEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RYGVEGERSMGMYTAVPNVDLMNKIFCVDVEHIDQFEIVRRSQLIPDERARKGRLMIE
ERAZANHYDGKYLTPEKLELQIKSYHAVREIVERMELDFVGIKGQLELTEHFVTMDVT
ETRINDPYDWEGEHEPIVCATFADSDGALTWQIPKLIAKTPVLFADVRHYVEEYDVLD
LCNGSNHATYFPAASEPDPERMKKVEFYPQTFFYFPAGGAAVKHIAAPGKNTKGRLTRE
DGRYRFTVVPGEFVDFGEEKNYEIADSIQNNWPHAFLKMETPIDEFLAKYSSNHIHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /product="alpha-L-fucosidase, putative"
/protein id="AAD35394."
/db_xref="G1:4980806"
/translation="MISMKPRYKPDWESLREHTVPKWPDKAKFGIFIHWGIYSVPGWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TPTGELGKVPMDAWFPQNPYAEWYENGLRIKESPTWEYHVKTYGENFEYEKFADLFTA
EKWPPQEWADLFKKAGAKYVIPTTKHHOGFCLWGTKYTDFWSVKRGFRAENLVGDLAKA
VREAGLRFQYYSGGLDMRFTTEPIRYPEDLSYIRPWTYEXADYAKQVMELVDLYLD
DVLWNDMGWPEKGKEDLKYLFAYYNKHPEGGSVNDRWGYPHWDFRTABYHVNYPGDLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GYKWEFTRGI GLSFGYNRNEGPEHMLSVEQLVYTLVDVVSKGGNLLLANVGPKGDGTI P
DLQKERLLGIGGEWLRKYGDAI YGTSVWERCCAKTEDGTEI RFTRKCNRI FVI FLGI PT
GEKI VI EDLNLSAGTVRHFLTGERLSFKNVGKNLEI TVPKKLLETDSI TLVLEAVEE"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EVIRAKEIVWKPSIAKREARRLAEEGAEVTIFNYAVWAFPHFTVLASKFAPGPFLLFS
NINPQYPGMVAMLSAAGALEQDGTKHYRWWGNIKDEKVLRKVWAFIRAGSTYKKLRGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8014. .9363
/gene="TM0306"
8014. .9363
/gene="sim0306"
/note="similar to GB:M29877 SP:P04066 PID:1335066
PID:178409 PID:182779 percent identity: 56.91; identified
by sequence similarity; putative"
                                                                                                                                                                                                                                                                           /note="similar to GP:3560020 percent identity: 55.11; identified by sequence similarity; putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="similar to GB:L42023 SP:P44779 PID:1005445 PID:1220691 PID:1204865 percent identity: 56.91; identified by sequence similarity; putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /GDYVEE1KTFCE1AS1DFVLMK
                                                                                                                                                                                                                                                                                                                                 /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               table=11
                                                                                                                                                              /gene="TM0305"
/894. .8017
                                                                                                                                                                                                                              5894. .8017
/gene="TM0305"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9360. .10781
/gene="TM0307"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /gene="TM0307"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
/transT_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /codon_start=1
/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9360. .10781
```

gene

CDS

gene

CDS

complement(10798. .13092)

gene

```
6013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6427
                                                                                                                                                                                                                                                                                                                                                                                                        6073
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6311 GATGGAAGAAACGCTGGTGAAAGATTGCAGGTAGATCCAAATTTCAGTAGCGTGCTTTTT 6370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGly 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerLeu 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             243 HisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGly 262
                                                                                                                                                                                                                                                                                                                                                                     45
                                                                                                                                                                                                                                                                                                                                                                                                                                               65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        85
                                                                                                                                                                                                                                                                                          25
/gene="TM0308"
complement(10798. .13092)
/gene="TM0308"
/note="similar to PID:1881669 percent identity: 63.77;
                                                                                                                                                                                                                                                                                                              GlyAlaThrTrpGlnIleThrProLeuPro-----PheLysLeuGlyGlyAsnMet
                                                                                                                                                                                                                                                                                                                                                                                        GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6194 GGAAAATACACTCGGGATTGG---GCAGGATATGGAGCCATTCTTATCTCGGAAGATTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6251 GGTGAAACGTGGACAATTGTAAACCTTGACAAGTATGGGATAAAAGTTGGTGGGAACGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyrThrTrpSerAsnValAla1leGlyGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6074 GAGGAAACAAAGAGGGGGAAGCAATTATTTTGAGTTTTTTGAGAAGAGAGTCCGGATTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlyGlnAlaSerLysThrilePheValGlyValAlaAspProAsnAsnProValPheTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhellePro
                                                                                                                                                                                                                                                                                                                                                                                                                                               46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --------AGTGTAACGTTTGTGCTATTCGATGAGAAAAGTGGAGAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6428 AAAGTGGATAGTTTTCCCTCAACC--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6551 ACTGAAGATGGTGGAACCACCTGGAATGTTCTTCCAAACCTGCCGAACGATCTCATTCCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::
                                                                                                                 13278
264
123
275
89
                                                                                                                                     Matches:
Conservative:
Mismatches:
                                                                                                                                                                                            Indels:
                                                                                                                   Length:
                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                     (1-13278)
                                                                                                                                                                                                                                                   US-09-917-376-3 (1-740) x AE001712
                                                                                                                 4.83e-50
1160.50
51.53
35.15$
28.75$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
```

8 8 8 8 8

_											-							
9919	322 6817	342	362 6910	382 6940	402	422 7045	436	456 7165	476	496	516 7324	533	553	570	588 7558	608	624	643 7735
6725 TACGATGTAACACCAATGAAAGGAGATTTCGGATATTGTGGA	303 LeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrp ::: ::: ::::	323 ProAspThr11e11ePheArgSerThrAspGlyGlyAlaThrTrpThrArg11eTrpAsp 	343 TrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAsplleSerAlaGluProTrp :::::::	363 LeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAsp:::	383 GlualametalaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThr ::: ::	403 LeuTyralaThrasnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaPro :::::	423 MetVallysGlyleuGluGluThralaValAsnAspLeuIle	437 SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis	457 AlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr ::: ::: 7166 GAAGTTTGGACACCCACCATCATCTATGTACAAACCCCTAAAATGGACGTCTCTT	477 SerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspPro	497 SerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe :::	517 GlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSer :::	534 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla	554 ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAla	571 GlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGly	589 ThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuPro	609 SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGlu 	625 GlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySer ::: 7676 GGAGATATCTGGTTAGCACTTCAATGGAATGGACTTTACAGATCAAAAGATGGTGGAATA
															•			ŀ

g

8 8

g &

8

엄

ò

8 8 8 8

8 8

Search completed: May 12, 2004, 09:25:37 Job time : 11278.8 secs

g

6 8 6 8 6

qq

ð

g

8

셤

Scoring table:

Sequence:

protein

Run on:

Total number

Searched:

Minimum DB Maximum DB

```
Acasa 418 Prokaryot
Continuation (5 of
Acts 3934 Mycobacte
Continuation (38 o
Continuation (38 o
Ada 30090 DNA encod
Acas 6562 Prokaryot
Continuation (31 o
                                                                                                                                                                                                                                         Aav35620 Human SHO
Abg77491 S. aurant
Ab129005 Drosophil
Ab129004 Drosophil
Ab111445 Drosophil
Ab111445 Drosophil
Ab111445 Drosophil
Aca1718 Prokaryot
Aca47178 Prokaryot
Continuation (38 o
Aca50513 Prokaryot
Continuation (38 o
Continuation (38 o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Continuation (39 o
Adc01011 Enterchae
Abz35486 Human gen
Aca20847 Prokaryot
Aad62312 Candida a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Aca40310 Prokaryot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile; cellulose; gene; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Abt 42169 Toxicity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Continuation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Continuation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /transl_except= (pos:2869,aa:Xaa)
/note= "Xaa is an unspecified residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nucleotide sequence of the avicelase AviIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ALIGNMENTS
                                                                                                                                                                                       ACA26562
AAI99682_30
AAI99683_30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACA40310
AAI99682_03
AAI99683_03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAI99682_38
                                                                                                                                                                                                                                                                                                                                                                                                                                          AA199683_37
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAD62307
AAI99682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABT42169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
1. 2869
4 tags a
/patial
/product = "Avill!"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAD62312
                                                                                                                                                                                                                                                                                                 ABL19628
                                                                                                                                                                                                                                                                                                                                   ABL19629
                                                                                                                                                                                                                                                                                                                                                                     ABL11444
                                                                                                                                                                                                                                                                                                                                                                                                                        ACA20513
                                                                                                                                                                                                                                                                                                                   ABL29004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABZ77632 standard; DNA; 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus.
5.9 7407
4.8 5295
4.8 10000
4.6 110000
4.6 110000
4.5 110000
4.5 110000
4.5 11679
4.4 110000
4.4 110000
                                                                                                                                                                        11679
3222
110000
110000
32367
67459
2855
                                                                                                                                                                                                                                                                                                                                                                                    7191
25165
4323
110000
110000
110000
25473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6615
110000
110000
                                                                                                                                                                                                                                                                                 WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-JUN-2003
                                                                  192
184
184
182
182
181
179
177
177
174
174
                                                                                                                                                                                                                                                                                               174
174
173
170.5
170.5
170.5
170.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                          169.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          169
169
168.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168
166
165.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABZ77632;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Key
 Command line parameters:
-MODEL=frame+ p2n.model -DBV=xlh
-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cqn2_1/USPTO_spool/US0917376/runat_11052004_114534_28231/app_query.fasta_1.1429
-Q=/cqn2_1/USPTO_spool/US0917376/runat_1EX=p2n.rng -MINNATCH=0.1 -LOOPCL=0
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINNATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAR=100 -THR_MN=0 -ALIGN=15
-MODE=LCCAL -OUTFYT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09917376_@CGN_1 1_758_@runat_11052004_114534_28231 -NCPU=6 -ICFU=3
-NO_MMAD -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOKE=100 -LONGLOG
-DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRRANS=1 -XGAPEXT=0 -LONGROG
-FGAPEXT=7 -YGAPEDP=10 -YGAPEXT=0.5 -FGAPOP=6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Trichoder
DNA encod
Trichoder
Trichoder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleotid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Abz77632 Nucleotid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Acidother
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Prokaryot
                                                                                                                                          (without alignments)
3062.577 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Add22922 Abv76941 PAdd42060 TAdd24893 IAdd42054 Aaf14988 TACA42281 IACA42281                                                                                                                                                                                                                              ........YIGTNGRGIVYGDIGGAPSG
                                                                                                                      May 12, 2004, 03:07:10 ; Search time 1026.48 Seconds
                   version 5.1.6
- 2004 Compugen Ltd.
                                                                                        nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                       hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                       3373863 segs, 2124099041 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                   0.5
7.0
7.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABZ77632
ADD22922
ABV76941
ADD42060
ADD24893
ADD42054
AAF14988
ACA42281
                                                                                                                                                                                                                              ATTQPYTWSNVAIGGGGFVD.
                                                                                                                                                                                                                                                                                 , Xgapext
, Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  geneseqn2001bs:*
geneseqn2002s:*
geneseqn2003as:*
geneseqn2003bs:*
                   GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 geneseqn2000s:*
geneseqn2001as:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      geneseqn2004s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               geneseqn1980s:*
geneseqn1990s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Geneseg 29Jan04:*
                                                                                                                                                                                                                                                                                                                                                                                                                                           seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                              US-09-917-376-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length
                                                                                                                                                                                                                                                                               Xgapop 10.0,
Ygapop 10.0,
Fgapop 6.0,
Delop 6.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2869
2869
2823
2849
3668
5698
11103
                                                                                                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0
100.0
51.1
37.3
28.2
21.8
21.8
9.7
                                                                                                                                                                                                                                                                                                                                                                                                         ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4036
4036
2063
1506
1139
878.5
392.5
                                                                                                                                                                                            Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
```

Jatabase

Result

```
The present sequence encodes a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                               useful
ies for
                                                                                                                                                          New thermostable AviIII peptide from Acidothermus cellulolyticus, usef for degradation of cellulose or in generating anti-AviIII antibodies f purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                 Himmel ME
                                              Adney WS, Vinzant TB,
                                                                                                                                                                                                                                                                           Claim 3; Page 24; 44pp; English
  (MIDE ) MIDWEST RES INST
                                                                                           2003-248177/24
                                                                                                                  P-PSDB; ABP73015
                                              Ding S,
                                                                                         WPI;
```

2869 740 0 0 0 Length:
Matches:
Conservative:
Mismatches:
Indels: Gaps: 9.08e-252 4036.00 100.00\$ 100.00\$ 100.00\$ Percent Similarity: Best Local Similarity: Alignment Scores: ouest Local S Query Match: DB:

TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100 CTGTATTTCGGCGCCCCCGAGCGGCTCTGGAGAAGCAATCCAGATTCCGGCGCGACC 618 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80 ATGTATCGATGGGATGCCGCCAACGGGCGGTGGATCCCTCTTCTGGATTGGGTGGATGG LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly AACAATTGGGGGGTACAACGGCGTCGTCAGCATTGCGGCAGACCCGATCAATACTAACAAG GTATGGGCCGCCGTCGGAATGTACACCAACAGCTGGGACCCAAACGACGGAGCGATTCTC CGCTCGTCTGATCAGGGCGCAACGTGGCAAATAACGCCCCTGCCGTTCAAGCTTGGCGGC AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAsnIle AACATGCCCGGGCGTGGAATGGCGGCGGCTTGCGGTGGATCCAAACAATGACAACATT AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAsp GGGATCGTCTTCAATGAAGGTGCACCGGGAATTCTGTACGTGCGGACGACATCGGGGGG MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly x ABZ77632 (1-2869) US-09-917-376-3 (1-740) _ 139 199 259 379 439 21 41 61 319 81 101 121 499 141 559 161

셤

셤

ò

a a

ò

셤

ò

ò

258

40

198

요

ò

유

à

g

ò 셤 ઠે

ð

20

318

9

378

Ор	619	TGGTCCCAGATGACGAACTTTCCGGACGTAGGCACGTACATTGCAAATCCCACTGACACG 678
රු දි	181	ThrGlyTyrglnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer 200
3 8	5 6	Accedentatendsocontations of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the contraction of the
h 원	739	TCGCTCGGGCAAGCGAGTAAGACCATTTTTGTGGGCGTGGCGGGATCCCAATAATCCGGTC 798
ζ	221	PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240
Dp	799	TTCTGGAGCAGAGACGGCGGCGGCGGCGGCGGGCGGGGCGCCGGCCGGCTTC 858
රු දි	241	IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
ì ò	261	ThrdlvGlvProTvrAspGlvSerSerGlvAspValTrpLvsPheSerValThrSerGlv 280
. AG	919	ACGGTGGTCCGTATGACGGGGGCCCCGCGCGTCTGGAATTCTCGGTGACCTCCGGG 978
č	281	ThrTtpThrArg11eSerProValProSerThrAspThrAlaAsnAspTyrPheG1yTyr 300
DP DP	979	ACATGGACGCGAATCAGCCCGGTACCTTCGACGGACACGGCCAACGACTACTTTGGTTAC 1038
<u>ک</u> دو	301	SerGlyLeuThrileAspArgGlnHisProAsnThrileMetValAlaThrGlnIleSer 320
3 ,	1037	אפריפיריו ראירואורטאיריפיריאפאראיריפיאאוראיריפיאאוראיריפיאיאוריפי די די די די די די די די די די די די די
දු පු	321	TrpTrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIle 340
ò	341	TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 360
qq	1159	TGGGATTGGACGAGTTATCCCAATCGAAGCTTGCGATATGTGCTTGACATTTCGGCGGAG 1218
δ,	361	ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp 380
OP	1219	CCTTGGCTGACCTTCGGCGTACAGCCGAATCCTCCGTACCCCAGTCCGAAGCTCGGCTGG 1278
&	381	MetAspGluhlametAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 400
qq	1279	ATGGATGAAGCGATGGCAATCGATCCGTTCAACTCTGATCGGATGCTCTACGGAACAGGC 1338
상 임 양	401	401 AlaThrLeuTyrAlaThrAsnAspLeuThrLygTrpAspSerGlyGlnIleHisIle 420
ò	421	AlaprometVallysGlyLeuGluGluThrAlaValAsnAspLeulleSerProProSer 440
ΩP	1399	GCGCCGATGGTCAAAGGATTGGAGGAGACGGCGGTAAACGATCTCATCAGCCCGCCGTT 1458
ò	441	GlyAlaProLeuileSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 460
Db	1459	GGCGCCCGCTCATCAGCGCTCTCGGAGACCTCGGCGGCTTCACCCCACGCCGACGTTACT 1518
ò	461	AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 480
Dp	1519	GCCGTGCCATCGACGATCTTCACGTCACCGGTTCACGACCGGCACCAGCGTCGACTAT 1578
à	481	AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro 500
Dp	1579	GCGGAATTGAATCCGTCGATCATCGTTCGCGCTGGAAGTTTCGATCCATCGAGCCAACCG 1638
ò	501	AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGlu 520
DP	1639	COACAGGCACGTCGCGTTCTCGACAGACGGCGGCAAGAACTGGTTCCAAGGCAGCGAA
۲۵ ·	N	540
ପ୍ର	1699	CCTGGCGGGGGGACGGCGCGCACCGTCGCCGATCGGCCGACGGCTCTCGTTTCGTC 1758

g S

ò 요 8 g ò g ò QQ Š 셤 8 유 ò g Š

```
New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTGTATTTCGGCGCCCCGAGCGCCAAAGGCTCTGGAGAAGCACACAGATTCCGGCGCGACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TrpSerGlnWetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyGlyPheValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AACATGCCCGGGGGTGGAATGGGCGGAGGGCTTGCGGTGGATCCAAACAATGACAATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlylleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly
                                                                                                                                                      The invention relates to an isolated polynucleotide molecule encoding thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucalotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. Method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents DNA encoding the Acidothermus cellulolyticus avicelase AviIII.
                                                                                                                                                                                                                                                                                                         BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                                                                                                                                                                                                              2869
740
0
0
0
                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
            Ä
                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
            Himmel
                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-3 (1-740) x ADD22922 (1-2869)
                                                                                                                                 Claim 17; SEQ ID NO 2; 29pp; English.
            TB,
                                                                                                                                                                                                                                                                                                                                                9.08e-252
4036.00
100.00%
100.00%
                                      2003-810853/76
            ws,
                                                                                                                                                                                                                                                                                                                                                            Score:
Percent Similarity:
Best Local Similarity:
Query Match:
                                                    P-PSDB; ADD22921.
             Adney
                                                                                                                                                                                                                                                                                                          Sequence 2869
                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
            ŝ
             Ding
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                      2058
                                                                                                                                                                                                                                                                                                          2118
                                                                                                                                  1879 ACTITCTATGCCCTATCCAATGGAACCTTCTATCGAAGCACGGCGGCGGCGTGACATTC 1938
                                                                                                                                                                                     AAGTCTGCGCCCCGGGTCGTCATACCCCAGCCGTCTTTGTCGTCGGCGCACGATCGGAGGCGTT 2178
                                                                                                                                                                                                                                                                                                                                                                                                     2358
                                                                                         1819 GCTGCTTCGCAAGGTGTTCCCGCCAATGCCCAGATCCGCTCAGACCGGGTGAATCCAAAG 1878
                                                                                                                                                                                                                                                                                                                                                                                                                                          720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                740
                                                                                                                                                                                                                            640
                                                                                                                                                                                                                                                                                099
                                                                                                                                                                                                                                                                                                                                     680
                                                                                                                                                                                                                                                                                                                                                                                       700
                                                                580
                                                                                                                      9
                                                                                                                                                                         620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   agricultural biomass; municipal solid waste;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TyrileGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                          2239 CACCAATACGGAAATTGGGGACAAGCAATCACCGGTGACCACGCGGATTACGGCGGGTG
            ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe
                                                                                                                                                                                                                          ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn
                                                                                                                                                                                                                                                      1999 CCTGGAAAAGAAGGCGATCTGTGGCTCGCTGCATCGAGCGGGCTTTACCACTCAACCAAT
                                                                                                                                                                                                                                                                                             ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2299 TACATAGGCACGAACGACGCCGTGGAATTGTATACGGGGAACATTGGTGGTGGTGGTGCGCCCTCCGGA
                                                                AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArqSerAspArgValAsnProLys
                                                                                                                                                                        GlnProvalAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal
                                                                                                                                                                                                                                                                                GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGly
                                                                                                                                                                                                                                                                                                                                     LysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyVal
                                                                                                                                                                                                                                                                                                                                                                                                                                            HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus avicelase AviIII DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    glycoside hydrolase; avicelase; ds; gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1. .2868
/*tag= a
/product= "Avill!"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cellulose reduction;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DING S.
ADNEY W S.
VINZANT T B.
HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-OCT-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                               2119
                                                                                                                                                                                                                                                                                                          2059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADD22922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADD22922;
                                                                561
                                                                                                                      581
                                                                                                                                                                                                                             621
                                                                                                                                                                                                                                                                                641
                                                                                                                                                                                                                                                                                                                                     661
                                                                                                                                                                                                                                                                                                                                                                                         681
                                                                                                                                                                                                                                                                                                                                                                                                                                             701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (DING/) I
(ADNE/) I
(VINZ/) I
(HIMM/) I
                                                                                                                                                                          601
```

원

198

20

The

318

80

9

100 438 120 498 140 558 160 618 180 678 200

258

g

g

ò

8 6 8 6

8 6 8 6

6 6 6

\$ B \$ B \$ B \$

9 9 9 8 6 8 6

8 8

g

Ś

8 6

ઠ

```
Xyloglucanase; enzyme; family 74; glycosyl hydrolase; cellulosic fiber; textile scouring; cellulose fiber processing; ratting; gene; ss.
                       GCTGCTTCGCAAGGTGTTCCCGGCCAATGCCCAGATCCGCTCAGACCGGGTGAATCCAAAG
                                                            CAACCGGTCGCGGCCGGTCTTCCGAGCAGCGGTGCCGTCGGTGGTTCCACGCGGTG
                                                                                                                                                                                                                                                     ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln
                                                                                                                                                                                                                                                                                                                                   ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe
                                                                                                                                           CCTGGAAAAGAAGCGATCTGTGGCTCGCTGCATCGAGCGGCTTTACCACTCAACCAAT
                                                                                                                                                                                   GGCGGCAGCAGTTGGTCTGCAATCACCGGGGTATCCTCCGCGGGGAACGTGGGGATTTGGT
                                                                                                                                                                                                              ACGGGGGCGTACCGCTCCGACGACGACGACGACCTGGGTACTGATGATGATGACCAG
                                                                                                                                                                                                                                                                                            HisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnLeuArgArgVal
                                                                                                                                                                                                                                                                                                        GlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal
                                                                                                                                  ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn
                                                                                                                                                                       GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "xyloglucanase enzyme'
                                                                                                                                                                                                                                                                                                                                                                                                                                                          sequence of a xyloglucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1. .2823
/*tag= a
/product= "xyloglucar
                                                                                                                                                                                                                                                                                                                                                                                               BP.
                                                                                                                                                                                                                                                                                                                                                                                               ABV76941 standard; DNA; 2823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-MAR-2002; 2002WO-DK000210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-MAR-2001; 2001DK-00000504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. .69
/*tag= b
70. .2820
/*tag= c
                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sjoholm C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (NOVO ) NOVOZYMES AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200277242-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                      03-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouffner F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nucleotide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gg
                                                   581
                                                                                         601
                                                                                                                                 621
                                                                                                                                                     1999
                                                                                                                                                                        641
                                                                                                                                                                                                              199
                                                                                                                                                                                                                                  2119
                                                                                                                                                                                                                                                      681
                                                                                                                                                                                                                                                                        2179
                                                                                                                                                                                                                                                                                            701
                                                                                                                                                                                                                                                                                                                                     721
                                                                                                                                                                                                                                                                                                                                                        2299
                                                                                                                                                                                                                                                                                                                                                                                                                  ABV76941;
            561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Jonesia
                                                                                                                                                                                                                                                                                                                                                                          8 8
                                                                                     6 g 6
                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                    셤
            8 8
                                                                                                                                                                     ò
                                                                                                                                                                                    8 &
                                                                                                                                                                                                                             8 B
                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                              CTTGGCTGACCTTCGGCGTACAGCCGAATCCTCCCGTACCCAACCTCGGAAGCTCGGCTGG 1278
                                                                                                                                                                                                                                                                                                                                                                                              1578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1818
                                                                                                                                                                                                                                                                        280
                                                                                                                                                                                                                                                                340
                                                                                                                                                                                                                                                                                                                                                                                     400
                                                                                                                                                                                                                                                                                                                                                                                                                        AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlyGlnIleHisIle 420
                                                         ThrGlyGlyBroTyrAspGlySerGlyAspValTrpLySpReSerValThrSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn
                                                                                                                 ATCCCGCACAAGGGCGTCTTTGACCCGGTCAACCACGTGCTCTATATTGCCACCAGCAAT
                                                                                                                                                                                                                                                            TrpTrpProAspThrilellePheArgSerThrAspGlyGlyAlaThrTrpThrArgile
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaProMetVallysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCGGCTATCAGGCGATATTCAAGGCGTCGGTCTGGGTCGCTTTCGACAAGTCTTCGTCA
                                                                                                                                                           Acederecerareacedeaecreceeceacererecaarrerecereacerecees
                                                                                                                                                                                                                                                                                                      TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu
                                                                                                                                                                                                                                                                                                                                             ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                    MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerieuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal
                                                                                                                    859
                                                                                                                                                            919
                                                                                                                                                                                                  979
                                                                                                                                                                                                                                          1039
                                                                                                                                                                                                                                                                                                                        1159
                                                                                                                                                                                                                                                                                                                                                               1219
                                                                                                                                                                                                                                                                                                                                                                                                       1279
                                                                                                                                                                                                                                                                                                                                                                                                                                              1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1639
                                      739
                                                                                                                                                                                                                                                              321
                                                                                                                                                                                                                                                                                                      341
                                                                                                                                                                                                                                                                                                                                            361
                                                                                                                                                                                                                                                                                                                                                                                                                          401
                   201
                                                          221
                                                                             799
                                                                                                  241
                                                                                                                                        261
                                                                                                                                                                               281
                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                                                                                                                   381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      521
```

2238

```
The present sequence encodes a xyloglucanase enzyme, belonging to family 74 of glycosyl hydrolases. The enzyme is isolated from Jonesia sp. DSM14140. The enzyme is useful in processes for machine treatment of fabrics. It is also useful in the textile industry for improving the properties of cellulosic fibers, yarn, woven or non-woven fabric, and in a textile scouring process step. The xyloglucanase enzyme is also useful in the cellulose fiber processing industry for ratting of fibers such as hemp, jute, flax and linen. It is useful for preventing binding of certain soils to the xyloglucan left on the cellulosic material
                                                                      New xyloglucanase enzyme belonging to family 74 of glycosyl hydrolases, and endogenous to a bacterium, useful in the textile industry for improving properties of cellulosic fibers, yarn or fabric.
                                                                                                                                                                         Claim 11; Page 68-72; 76pp; English
WPI; 2003-092855/08
P-PSDB; ABB99489.
```

유

ò

⋧

6 6 6

ò

Othe
0
'n
0
Ë
591
Ö
764
່ວ
803
À,
665
BP;
2823
Sequence

189 189 189 189 189		105	125	145	165	184	204	1 224
heasn rrraac rpasp rraar	Alaval SCAGCA	AspGln ::: SACAAA	31yArc 3GTCGC	31yAla 3GTGCG	MetThr ::: GTGACG	-TyrGln GTACACC	31yGlr 3GrcAQ	SerArg
es: 3.88e-124 Length: 2823 2063.00 Matches: 386 2121\$ Conservative: 112 ilarity: 67.21\$ Mismatches: 225 ilarity: 52.09\$ Mismatches: 125 Indels: 9 3 (1-740) x ABV76941 (1-2823) TYTHATADSEASANVALAIAILEGLYGLYGLYGLYPHeVALASPGLYILEVALPHEASN ::	AsnGlyValValSerIleAlaAlaAsproIleAsnThrAsnLysValTrpAlaAlaVal TccGAAATCTGTCCTTGGCACTGACCTGTTGACACCAACGGGTGTACCTTGCACA	GlymettyrthrasnSerTrpaspProasnaspGlyala11eLeuargSerSerAspGln 	GlyalathrtrpGln1lethrProLeuProPheLysLeuGlyGlyasnMetProGlyarg 	GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 	ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 	; §	SerAspIleGlnGlyvalValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln 	AlaSerLygThrilePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg
3 spGlyI CTGGAA 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT 1\metT	rsval1 :: : GCGTGT	euArgS rgcGTT	lyasna 3CAACA		hrTrps ccreec	hrThrd ::: CTTCTC	erSerS ::: AGCGA	rovalPhe'
2823 386 112 225 18 19 19 11 TGTGCC TGTGG1 1G1\TT	rAsnLy : CAACCC	allete TATTT	uG1yG] : GGGAG(pAsnI : : TAAAG	yalari 	raspti -Gatg	BSerSe ::: CACAT	nAsnP1
tive: les: GlyPh -	eAsnTh :::: TGACAC	GlyAl GGTGC	LysLeuC ::::: :cgrGrG	AsnAs :: AATAA	Sergl	ProTh	ASPLY GACCC	ProAs
es: 3.88e-124 Length: 2063.00 Antches: 386 2063.00 Antches: 318 36.7.21\$ Mismatches: 112 Allantity: 57.01\$ Gaps: 3 (1-740) x ABV76941 (1-2823) TYTHTTPSETASNValAITIEGIYGIYGIYPHeVAlAspGIYII THISTILL!	Prolle	Asnasr ::: CAAAAC	ProPhe	Proasr CCAAAC	Thrast ACTGAC	ASIPHEProASPValGlyThrTyrIleAlaAsnProThrAspThrThrGly ::: :: AGCTTCCCCAATGCTGGTAACTACGTCGCCGATGCTTCTGGT	AlaPhe ACTTTT	AlaAst
Lenc Cons Matc Cons Mind Gaps (1-2823) 116919G 	laAsp CTGAC	SpProj ATCCA	roleu 	alAspl TTGAC	rgser :: AGTCA	AsnPheProAspValGlyThrTyrIleAla ::: ::: AGCTTCCCAATGCTGGTAACTACGTCGCC	rpvali GGGTA	lyval
24 (1941 (11Alai (17GAGA -ELEUT :::::	eAlaA :	rTrpA crggg	eThrP GACGA	euAlaV rcgccg	euTrpA : :ATGGA	YThrT	1ValT TGTGT	eValG
3.88e-124 2063.00 52.03.00 67.21\$ 51.11\$ x ABV76941 werasnvalal CGAACGTTGA rcgly1leLe rcgly1leLe rcgly1leLe rcgly1leLe rcgly1leLe rcgly1leLe	lserile : ::: Grccrrg	rAsnSe : CATGG	pGlnI] ::: GGAAA	uArgle ACGCC	sGlyLe TGGACT	pva1G] : TGCTG(nG1yVa AGGTG1	hrilePheval
3.4 200 200 1149: 52 7 7 7 7 7 7 7 1740) x J 1741	ValVal :::::: ATTCTG	Tyrth :: TACTO	ThrTr	GlyGl GGAGA	GlyLy GGTAA	ProAs :: CCCAA	11eg1 	SerLysTh
inlarity: Similarity: 1: 176-3 (1-7 6 TyTThr 6 TyTThr 7 (1-1: 130 TATAGC 1: 190 CAGAAA 46 AlaAla 150 GACTCC	AsnGly ::: rccGGA	31yMet 3GCACG	31yAla 3GCGAG	31yMet 3GGATG	ProSer 3AATCG	AsnPhe GCTTC	SerAsp 3GGCAA	AlaSer
11 in in in in in in in in in in in in in		370	106 (126 0	146 F	166 4	185 8	205 7
ument . No. . No. . Loca . Mat . Mat								
Aliginary Prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the prediction of the pr	<i>장</i> 음	ò a	ۇ <u>م</u>	<u>ک</u> و	6 G	<u>ک</u> ۾	<u>ک</u> ۾	ò

ઠે

1786

ઠે

```
1437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TACTCAGGGCTGGCAGACATCGGTGGCTTCACCCCACAAGATATCTCTCAGGTGCCTAAC 1497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1731
                                                                                                                                                                                                                                                                                                                                       GGTTCACAGCCGCGTACAAGCCAGTACAACCACGATTACTCAGGGGCTCCCTGGCTTGAC 1197
                                                                                                                                                                                                                                                                                                                                                                            ||||||||
TTTGGCAATACAGCCAAAGAGCCTGAAGCTAACCCCAAACTCGGGTGGATGACCCAGTCC 1257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGCGGAACGAACCTCACAAACTGGGATAAGGGCAAAAAAGTTGATATCACTGTCAAGGCC 1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   563
780
                       244
                                               840
                                                                     264
                                                                                             900
                                                                                                                    284
                                                                                                                                           960
                                                                                                                                                                   IleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr 304
                                                                                                                                                                                                                  324
                                                                                                                                                                                                                                                               ThrileilePheArgSerThrAspGlyGlyAlaThrTrpThrArglleTrpAspTrpThr 344
                                                                                                                                                                                                                                                                                                                SerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThr 364
                                                                                                                                                                                                                                                                                                                                                                384
                                                                                                                                                                                                                                                                                                                                                                                                              404
                                                                                                                                                                                                                                                                                                                                                                                                                                                             424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       523
                                                                                                                                                                                                                                                                                ATGTACGTATATCGCTCCACCCACCGGGGGAAAACGTGGTCGCCTATCTGGGAAACTGAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1558 AAACCGGCAACTGTGCGTGCAGGCAAA-----TCTATTTCTGGGGAAACCACCTCA
ACAACACAAACGATCTACGTAGGTGTAGCAGACAAGCAGAACAACGTTTATCGATCTACC
                                    TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArg
                                                                                                                                   IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp
                                                                                                                                                                                                                                PheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAla
                                                                                                                                                                                                                                                                                                                                                                                                              MetalaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Thr --- IlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnProSerIleIleValArgAlaGlySerPheAspProSerGlnProAsnAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----gaaddigccgcrccrcGacGrrccaccaacaGrddcrcrccrcdidgrcaacGGrG
                                                                      GlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaThrAsnAspieuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyr
                       AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLys
 721
                       225
                                              781
                                                                      245
                                                                                             841
                                                                                                                                                                     285
                                                                                                                                                                                                                    305
                                                                                                                                                                                                                                           1018
                                                                                                                                                                                                                                                                  325
                                                                                                                                                                                                                                                                                          1078
                                                                                                                                                                                                                                                                                                                  345
                                                                                                                                                                                                                                                                                                                                                                365
                                                                                                                                                                                                                                                                                                                                                                                       1198
                                                                                                                                                                                                                                                                                                                                                                                                               385
                                                                                                                                                                                                                                                                                                                                                                                                                                       1258
                                                                                                                                                                                                                                                                                                                                                                                                                                                              405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      544
                                                                                                                      265
                                                                                                                                             901
                                                                                                                                                                                             961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465
                      8 6
                                                                                                                    8 6
                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                   Ωp
                                                                                                                                                                                                                                                                                                                                 유
장
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         요 장
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
119 TTTTCATGGAAGAACGTCAAGCTCGGCGGCGGCGGCTTCGTCCCCGGCATCATCTTC 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspileGlyGlyMetTyrArgTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45 AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-3 (1-740) x ADD42060 (1-2849)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
          intron
                                      exon
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                CGTCAAGGTCATCTCTGGGCTTGCTGGAGTATCGGGAAGCACCTATGGGATGTGGCGG 2019
                                                                                                                                ||||||||||||||::
|GTTTCGGTAAAGCAACCAGTAGTAGCGGATACCCAGTGATTTTCACATCCGCCAAGATT 2139
                                                                                                                                                                                                                                               |||::: :::|||:::
|GACAACCACAGTGGTATTGGACTGGCGCATCAATCACGGTGACCCAGATGTCGAA 2259
                                                                                                                                                                                                                                                                                          2260 GGGCGCGCTCTACATTGGTACCAACGGTCGCGCATTATCGTTGGTGATTCCTCCACTCCC 2319
                                                                                                                                                                                                                 2140 GATGGAGTTCGAGGAATTTTCCGGTCTGACGATGAAGGCAAGACCTGGAAACGCATCAAC 2199
                                                                                                                       657
                                                                                LysGluGlyAspLeuTrpLeuAla------AlaSerSerGlyLeuTyrHis 637
                                                                                                                                                          GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrile 677
                                                                                                                                                                                              GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsn 697
                                                                                                                                                                                                                                    AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnLeu 717
                                                                                                                                                                                                                                                                         ArgArgValTyrileGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAla 737
                                                                                                                                                                                                                                                                                                                                                                                                                                             Trichoderma reesei; Family 74 xyloglucanase; beta-1,4-glycosidic linkage; xyloglucan hydrolysis; biomass degradation; cellulose; hemicellulose; ethanol production; detergent composition; fabric treatment; textile treatment; gene; ds.
       584 AlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProVal
                                            Ala---AlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGly
                                                                                                                     SerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                           Trichoderma reesei Family 74 xyloglucanase gene, SEQ ID NO:1 #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           74 xyloglucanase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Contains 3 introns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Family
                                                                                                                                                                                                                                                                                                                                                                     BP.
                                                                                                                                                                                                                                                                                                                                                                     DNA; 2849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1228. .1277
/*tag= e
/number= 2
1278. .2664
/*tag= f
                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441. .1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          product=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  '*tag= c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              '*tag= d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .2743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "C
59. .381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                number=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'number=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'number=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      *tag=
                                                                                                                                                                                                                                                                                                                                                                    ADD42060 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hypocrea jecorina.
                                                                                                                                                                                                                                                                                                                                CCG 2322
                                                                                                                                                                                                                                                                                                             Pro 738
                                                                                                                                                                                                                                                                                                                                                                                                          15-JAN-2004
                                                                                623
                                                                                                                                                                                               678
                                                             1906
                                                                                                   0961
                                                                                                                                        2020
                                                                                                                                                                            2080
                                                                                                                                                                                                                                                       2200
                                                                                                                                                                                                                                                                                                                                2320
                                           604
                                                                                                                     638
                                                                                                                                                          658
                                                                                                                                                                                                                                                                         718
                                                                                                                                                                                                                                                                                                                                                                                       ADD42060;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         intron
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  intron
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    exon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               exon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             exon
                                                                                                                                                                                                                                                                                                                                                            ADD42060
                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                    셤
                                      8
                                                       셤
                                                                               8
                                                                                             용
                                                                                                             જે
                                                                                                                                   g
                                                                                                                                                         ò
                                                                                                                                                                      g
                                                                                                                                                                                         8
                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                            ò
       à
```

9

44

TyrThrTrpSerAsnValAlaIle---GlyGlyGlyGlyPheValAspGlyIleValPhe

2849 329 121 260 87 23

Length: Matches: Conservative: Mismatches:

4.23e-88 1506.00 56.60% 41.38%

Indels: Gaps:

```
The invention relates to a Family 74 xyloglucanase (ADD42055) from the fungus Trichoderma reesei (Hypocrea jecorina), and nucleic acids encoding it (e.g., ADD42064). The enzyme hydrolyses the beta-1,4-glycosidic linkages in the backbone of xyloglucan to xyloglucan oligosaccharides. The invention also relates polypeptide sequences at least 70% identical comparison of the enzyme, expression vectors and host cells comprising a nucleic acid of the invention, the recombinant production of the enzyme, and mutant enzymes and the nucleic acids encoding them. The xyloglucanase of the invention can be used in the degradation of cellulose- and hemicallulose-containing biomass to produce ethanol. It can also be used in a detergent composition for treating fabric during a machine washing cycle. The present sequence represents the gene encoding Trichoderma reese is family 74 xyloglucanase. Note: A comparison of this sequence with that the present sequence listing (also referred to as SEQ ID NO:1) shows that the present sequence listing (also referred to as SEQ ID NO:1) contains alternate 60 bp blocks of sense strand and complementary strand
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New polypeptides having Family 74 xyloglucanase activity, and encoding nucleic acid molecules, useful for degrading cellulose- and hemicellulose containing biomass to ethanol or as a detergent.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP; 594 A; 882 C; 822 G; 551 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               <u>ن</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (NOVO ) NOVOZYMES BIOTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 1; Fig 1; 96pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-APR-2002; 2002US-0373987P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-APR-2003; 2003WO-US011831
/number= 3
2665. .2723
/*tag= g
/number= 3
                                                                                                                                                               2724. .2743
/*tag= h
                                                                                                                                                                                                                                                    number= 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zaretsky E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2003-845528/78.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-PSDB; ADD42055
                                                                                                                                                                                                                                                                                                                                    WO2003089598-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2849
                                                                                                                                                                                                                                                                                                                                                                                                               30-OCT-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Michael R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sednence.
```

6 6 6 6

```
셤
                                                                                              셤
                                                                                                                                               g
                                                                                                                                                                          8
                                                                                                                                                                                                   용
                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                       Š
                                                                         8
                                                                                                                            δ
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                슝
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention relates to polypeptide and polynucleotide sequences from Xanthomonas campestris which may be used for activity reduction or enhancement using directed genetic engineering. A transformed cell or organism having reduced or enhanced activity of at least one such protein e.g. galactomannanase can be generated by disrupting the gene encoding the protein. The activity of the protein is reduced by the presence of an antisense nucleic acid sequence. The nucleic acid sequence of the gene encoding the protein is a recombinant sequence having at least one mutation as compared to the wild-type gene encoding the protein. The transgenic cell or microorganism are useful for producing xanthan gum, which are useful for providing formulations and properties, such as longtern suspension and emulsion stability in alkaline, acid, and salt solutions, temperature resistance, and pseudoplasticity. The present
                                                                                                                                                                                                                                                                                                                                                                                                                                  transformed cell or organism having reduced or enhanced activity of least one protein, useful for producing xanthan gum, which are useful
                                                                                                                        Directed genetic engineering; galactomannanase; reduced activity; enhanced activity; xanthan gum production; suspension stability; emperature resistance; pseudoplasticity; amylase; cellulase; extracellular protease; intracellular protease; glucose dehydrogenase; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BP; 570 A; 1181 C; 1292 G; 625 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequence encodes an enzyme relating to the present invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3668
267
120
295
79
                                                                                                  encoding Xanthomonas campestris cellulase #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            formulations and properties
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 15; SEQ ID NO 21; 135pp; English.
                        BP
                       ADD24893 standard; DNA; 3668
                                                                                                                                                                                                                                                                                                      28-MAR-2001; 2001US-0279493P
                                                                                                                                                                                                                                                                              10-AUG-2001; 2001US-00927827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.13e-64
1139.00
50.85%
35.09%
28.22%
                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                     Ramseier TM;
                                                                                                                                                                                                     Kanthomonas campestris
                                                                                                                                                                                                                                                                                                                                            RAMSEIER T M.
                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2003-625389/59.
P-PSDB; ADD24919.
                                                                                                                                                                                                                                                                                                                                BOWER S
                                                                                                                                                                                                                            US2003036176-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                         providing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 3668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gnment Scores:
                                                                                                                                                                                                                                                                                                                                                                     SG,
                                                 ADD24893;
                                                                                                                                                                                                                                                                                                                                                                                                                                                least
                                                                                                                                                                                                                                                                                                                                            (RAMS/)
                                                                                                                                                                                                                                                                                                                                (BOWE/)
                                                                                                                                                                                                                                                                                                                                                                     Bower
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                    New
RESULT
```

GlylleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly 40

AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAsp 20

(1-3668)

x ADD24893

JS-09-917-376-3 (1-740)

Н

용 Š

21

ò

Percent Similarity: Best Local Similarity:

Matches: Conservative: Mismatches: Indels:

```
CGCACGCCGCGCGACGACGTGTATCGCAGCGTGGATGGTGGCAAGCACTGGACGCCGTTG 2221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----TTGGCCGATGCGGTGTTCGATCACAGCGCCGCG 2254
                                                                                                                                                                                                                                                                                                                                                                                              140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MetAspGlualaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 400
                                                                   ||||||| :::::||| |||:::::
                                                                                                                                                                                                                                                                                                                                                                                                                       1574 AACCAGCTGGGCCCAATGGCGAGCGGCTGGCGTGGACCCGCACGATGGCCGCGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TATGTGTCCGAAGATGCCGGCCGCAGTTGGGCACCGGTGGCCGGGCCAACCACGCGGGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2045 CGCTGGCGTGAGATCACCCGATTCCGCAG----CCAGCCAGTGGCGATGGATTCGGCTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGTGCGGTGGGTGCGCAACATCCGCAGGTGCTGCTGGCCAGCACCTTCCGCCGT
                                                                                                                              AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                         GACGACTGGAACCTGATGGGCATCGACGCATTCGCCGTCGACCCCGCCGATGCCGATGCG
                                                                                                                                                                                                                                                                                                      ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LeuTyrPheGlyAlaProSerGlyLySGlyLeuTrpArgSerThrAspSerGlyAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 201 SerLeuGlyGlnAlaSerLysThrilePheValGlyValAlaAspProAsnAsnProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          221 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         281 ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp
                                                                                                                                                                                                                    ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu
                                                                                                                                                                                                                                                                                                                                                                                              AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAspAsnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp
GGTGTGCTGTTTCATCCCGCCGAACGTGGTCTGGCCTATGCGCGCACCGATGTGGG
1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2102
                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
US-09-917-376-3 (1-740) x ADD42054 (1-5698)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GGTGTGATGCGCTGGCAGGACGGT-----GCGCTGCAGGTGCTCTCACAGCCGGACGA 3106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2789 GCGCACGTCCGACTTCGGCGCAGTGGCAGCGTGTGAACGCCCTGCCGAACCCCGGGTG 2848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGTGGTGACCGATCGCGTGGACGCACGCGCTGGTATGCGGTGGATGTCGCCAGCGGCA 2908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2909 GCTGTACGAGAĞCACCGATGCCGCGCGCGGTTTCCGT-----ĞCĞACCGĞTGTACAGGT 2962
ATGGGG---GCGCTGGCGATCGATCGTTCGACGCCAACCATGCGCTGTTCGTGACCGGC 2341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CAAGCAATGGACTGCGTTTGCGAGTGAGCCGCCAGCCGGGCGGCGCGGCGCGCATTGC 2734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2963 GGGCAGCCCGGCGCGCGGTGAGCGCACCCGGCCGCAGCTGCGTCCGGACCCGTGGCGCC 3022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               733
                                                                                                                                                                                                                                                                                                                                                                                                                                     472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pValleulleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAs 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nlleArgSerAspArgValAsnProLysThrPheTyrAla-----LeuSerAsnGlyTh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TACGGCATCTGGGCCTCGCGCAATCTG------CAGGATTTCGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAs
                                                               401 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle
                                                                                                                                                                                  ----MetValLysGlyLeuGluGluThrAlaVal
                                                                                                                                                                                                                                                                                                               433 AsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGly
                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Accerecececececes Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces Acces 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 yLysAsnTrpPheGln---GlySerGluProGlyGlyValThrThrGlyGlyThrValAl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        aAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----ProSerSerGlyAlaVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1GlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            673 lValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTr
                                                                                                                                                                                                                                               GCACCGCAGCGCCCCCCTGCAGTGGTTCCAGGACCGTGGGCTGGAAGAAACCGTGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Thr Thr Gly Thr Ser Val Asp Tyr Ala Glu Leu Asn Pro Ser Ile Ile Val Arg Ala Gly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerPhe-AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 551 lTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rPhelyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeu----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             rGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3023 AGGCGTGGTGTATCTGGCCAGCCCCGGCAAG----
                                                                                                                                                                                         AlaPro
                                                                                                                            2342
                                                                                                                                                                                  421
                                                                                                                                                                                                                                               2381
                                                                                                                                                                                                                                                                                                                                                                        2441
                                                                                                                                                                                                                                                                                                                                                                                                                                     453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3167
   2285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                         ò
                                                                                                            g
                                                                                                                                                                            õ
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
The invention relates to a Family 74 xyloglucanase (ADD42055) from the fungus Trichoderma reesei (Hypocrea jecorina), and nucleic acids encoding it (e.g., ADD42054). The enzyme hydrolyses the beta-1,4-glycosidic linkages in the backbone of xyloglucan to xyloglucan oligosaccharides.

The invention also relates polypeptide sequences at least 70% identical to the enzyme, expression vectors and host calls comprising a nucleic acid of the invention, the recombinant production of the enzyme, and mutant enzymes and the nucleic acids encoding them. The xyloglucanase of the invention can be used in the degradation of cellulose- and hemicallulose- containing biomass to produce ethanol. It can also be used in a detergent composition for treating fabric during a machine washing cycle. The present sequence represents a Trichoderma reesei Family 74 xyloglucanase-encoding DNA. Note: The present sequence (referred to as SEQ ID NO:1) is given as shown in figure 1 (also referred to as SEQ ID this sequence with that shown in figure 1 (also referred to as SEQ ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New polypeptides having Family 74 xyloglucanase activity, and encoding nucleic acid molecules, useful for degrading cellulose- and hemicellulose containing biomass to ethanol or as a detergent.
                                                                                                                                                                                                                                           Trichoderma reesei; Family 74 xyloglucanase; beta-1,4-glycosidic linkage; xyloglucan hydrolysis; biomass degradation; cellulose; hemicellulose; ethanol production; detergent composition; fabric treatment; textile treatment; ds.
                                                                                                                                                                                                                Trichoderma reesei Family 74 xyloglucanase-encoding DNA, SEQ ID NO:1 #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NO:1) shows that the present sequence is twice as long as the figure 1 sequence (ADD42060) and contains alternate 60 bp blocks of sense strand
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Other
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 5698 BP; 1145 A; 1704 C; 1704 G; 1145 T; 0 U; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5698
334
119
270
793
52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; SEQ ID NO 1; 96pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complementary strand sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Haas J;
                                                                                                         BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (NOVO ) NOVOZYMES BIOTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-APR-2002; 2002US-0373987P.
                                                                                                         DNA; 5698
                                                                                                                                                                                                                                                                                                                                                                                                                                                  17-APR-2003; 2003WO-US011831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.69e-47
878.50
29.92%
22.06%
21.77%
                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zaretsky E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2003-845528/78.
                                                                                                         standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                           Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB, ADD42055.
                                                                                                                                                                                                                                                                                                                                                                              WO2003089598-A2
                                   3344 T 3344
p 733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
Pred. No.:
                                                                                                                                                                             15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                 30-OCT-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Michael R,
                                                                                                                                            ADD42054;
 733
                                                                                                         ADD42054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                          ADD4205
                                                                      RESULT
                                                                                                         셤
```

0y 166 AsnPheProAspValGlyThrTyrIlealaAsnProThrAspThrThrGlyTyrGln	2	Oy 225 AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHis	Qy 248 Db 1801 Qy 264 Db 1861	265 TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArg	25 294	1000 308 GlinHisProAsnThrileMetValAlaThrGlinIleSerTrpTrpBroAspThr	328 2281 348 2341 362
7 ThrTrpSerAsnValAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsn ::: :::	45		86 GlyMetTyrThrasnSerTrpasp	94			1141 CCTCGGCCTCTCGCAGACCGACAGCTAGGCGGTTGAGGTGTAGATGAAACCACGA 147SerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet

	Qy 550 ValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsn 569 ::	3777 572 572	590 Phe		Qy 610 SerGlyAlaValGlyValMet	Qy 617	Qy 628 TrpLeuAlaAlaSerSer	4		645TrpSerAlalleThrGlyValSerSerAlaValAs	4365 TGGAACCTGTATGCCTGGTTGTGGATGGTCTAGCGGGACCCACACCCGAGTCCGAGCTTG	Oy 669 ProAlaValPheValGlyThrIleGlyGlyValThrGlyAla1YzArgser 686 	Qy 687 AspaspCysGlyThrTrpVal	Oy 695 704 Db 4539 CCTCTGTCGCCGCGGAGACTCCTCGAACATCCCAGGCTTCCAGGCTTCGGCTCCATCGAC 4598
		Oy 38711eAspProPheAsnSerAspArgMetLeuTyrdlyThrGlyAlaThrLeuTyrala 405	Db 2760 cagatrdacccaacadcaaccacrdacrtracdcaacadcaacadcaarcaaraaca 2819 Qy 406 Thr	2820 GTCTAACTGGGTTGGCTGTGGTGACCGAGATGCCGTGGCCTTACTGTTAGAAACCG 407ASnAspLeuThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLys :::	2880 CGCCACGATCTCACCAACTGGGGACACGCCCACAATGTGTCAATCCAATCAACAGACAG	Db 2940 GCGGTGCTAGAGTGACCCTGTGCGCGGTGTTACACAGTTAGGTAGTGACCGTCTG 2999 Qy 427LeuGluGluThrAlasnAspLeuIleSerProProSerGlyAlaProLeuIle 445 Db 2000 CGCARCGAGAAATTCTCGACCAGAAACTCACCAGAGAAAACAAGCAAG	446 Ser	447AlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaVal	Oy 462	463 ProSerThr1lePheThrSerProValPheThrThrGlyThrSerValAspTyrAla 3240 GGGGAGACGGTCTGGGCAACGCCCACATGGGCCACCTCGACGACGTCGACGTCGACGTCGACGGG	Qy 482Glu 482 Db 3300 AGCGTCTGCCAGACCCGTTGCGGGTGTACCCGGTGGAGCTGCTCGCAGCTGATCCGGCCC 3359	Qy 483 LeuAsnProSerIleIleValArgAlaGlySer	494PheAspProSerSerGlnProAsnAspArgHisValAlaPhe	508 SerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGly

113 F

```
The present in the present in describes a metrous fungal (FF) cell relative to expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes common for a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential common common feel of the microorganisms to be improved. New genes may be discovered, to of the microorganisms to be improved. New genes may be discovered, confittions, environmental stress, spore morphogenesis, recombination, conditions, environmental stress, spore morphogenesis, recombination, conditions, environmental stress, spore morphogenesis, recombination, advantages over genomic or random cDNA clones including elimination of advantages over genomic or random cDNA clones including elimination of conduction of the microarrays begues on function of the gene correspondent to facilitate analysis of the results. AAF07478 to AAF11247 represents
                                AGCACCAAGGTCGCCGGCAGCGGCCTGTAGGTCCCGAGGGTCCCGAAGCCGAGGTAGCTG 4658
                                                                                       724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cells
                                                                                                                                                                                                                                                                                                                                                                                                                expressed sequence tag; Fusarium venenatum; Aspergillus niger; Aspergillus oryzae; Trichoderma reesei; identification; recombination; culture condition; environmental stress; spore morphogenesis; metabolic pathway engineering; catabolic pathway engineering; ss.
                                                                  AsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention describes a method for monitoring differential
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     differential expression of genes in filamentous fungal ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Olsen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Clausen IG,
                                                                                                                                                          AsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaPro 738
                                                                                                                                                                                                                                                                                                                                                                                                 Multiple gene expression; filamentous fungal cell; EST;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kauppinen
                                                                                                                                                                                                                                                                                                                                                             Prichoderma reesei EST SEQ ID NO:7511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89; Page 3034; 3161pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         substrate of expressed sequence tags
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Shuster JR,
                                                                                                                                                                                                                                         AAF14988 standard; cDNA; 1103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BIOTECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-00273623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000WO-US007781
                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       fluorescence-labeled
                                                                                                                                                                                                                                                                                                               (revised)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (NOVO ) NOVO NORDISK
(NOVO ) NOVO NORDISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-594572/56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rey MW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200056762-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22-MAR-1999;
                                                                                                                                                                                                                                                                                                             15-SEP-2003
13-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RW,
                                1599
                                                                                                  4659
                                                                705
                                                                                                                                      725
                                                                                                                                                                      4713
                                                                                                                                                                                                                                                                             AAF14988;
704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Berka
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim
                                                                                                                                                                                                                       8
                                                                                                                                                               유
ò
                                                            ठे
                                                                                                                                    ઠે
```

```
341
                                                                                                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                     302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GlyGlyGlnIleHisIleAlaProMetValLys-GlyLeuGluGluThrAlaV 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----AlaProLeulleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspV 459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    528
                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             117 TCAAGGACCTGGCCTTTTCACCCGGGGGAAGCGAGCTTTTGGCCGCAAGTCCGGAGACG 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CAACAAGGIGGCCAITITITICCGAACGGCGCGCCCGACGINGGAA-----GCAAITC 686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     787
    ESTs from
                                                                                                                                                                                                                                                        303 TTÁCGGCÁCCGGÁATGACAATNTTTGGCGGCCACGÁTTTCACCAACTGGGACACGCGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |----ACATTGCCGCAGACGGTTTGGGCAACTCCCACATGGGCCACCTCGACGACGACGTC
                                                                                                                                                                                                                                                                                                                         AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro
                                                                                                                                                                                                                                                                                                                                                                                                               ---Pro
                                                                                                                                                                                                                                                                                                                                                                                                                                   TGGATCAAGAACAACTTTATCGATGTGACGAGCGAGTCACCGGTCCGATGGTCTNATCAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                erThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        528 lyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyG
                                                                                                                                                                                                                                        GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp
                                                                                                                                                                                                                                                                                               TrpProAspThr1le1lePheArgSerThrAspGlyGlyAlaThrTrpThrArg1leTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          477 ANCAACGGGTTTACCTTTTGCCCAGCAGAAACGACCTTGGG---------
                               the present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -------PheAspProSerSerGlnProAspArgHisValAlaPheS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GAACTAACGCT------GGNTCCGAACACGTTTCCATTGAAACGGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr-SerVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     479 AspTyrAla -- GluLeuAsnProSerIleIleValArgAlaGlySer------
ESTs from Aspergillus niger; AAF11854 to AAF14878 represents Aspergillus oryzae; and AAF14879 to AAF15377 represents ESTs Trichoderma reesei, which are all specifically claimed in the invention. (Updated on 15-SEP-2003 to standardise OS field)
                                                                      C; 307 G; 223 T; 0 U; 21 Other;
                                                                                                                                                                                                                                                                                                                                                                                                           TrpLeuThrPheGly---ValGlnProAsnProProValProSer--
                                                                                                               1103
1114
36
125
58
                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                     Indels:
                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alAsnAspLeuIleSerProProSerGly
                                                                                                                                                                                                            US-09-917-376-3 (1-740) x AAF14988 (1-1103)
                                                                      BP; 230 A; 322
                                                                                                             1.54e-16
392.50
45.05%
34.23%
9.72%
                                                                                                                                                        Similarity:
                                                                                                                                          Percent Similarity:
Best Local Similarit
                                                                      Sequence 1103
                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                      183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                   m
                                                                                                                                                                                                                                                                                                                                                     342
                                                                                                                                                                                                                                                                                               322
                                                                                                                                                                                                                                                                                                                                                                                                           362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       397
                                                                                                                                                                                                                                        302
                                                                                                                                                                                                                                                                                                                          63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              687
                                                                                                                                                                    Query Match:
DB:
                                                                                                                 ..
9
 8X333X8
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                        ሯ
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

```
proliferation (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway trequired for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
                                                                     839
                                                                                                                                                                                                 840 CGGGCGCCGTCATCGNCTCGGACAAGAACCAACAGCGTNTTCTACGCCGCTCCGGAT 899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for
                                      548 lnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProA
                                                                                                                                   laAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zyskind JW;
Xu HH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               prokaryotic essential gene; cell proliferation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ohlsen KL,
Forsyth RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 14; SEQ ID NO 30151; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                     900 cgácciririacercaccaagacaccece 928
                                                                                                                                                                                                                                                                   588 lyThrPheTyrArgSerThrAspGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Prokaryotic essential gene #23938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Malone C,
Carr GJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-0CT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACA42281 standard; DNA; 7407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-MAR-2002; 2002WO-US009107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-MAR-2002; 2002US-0362699P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ELITRA PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 drug design, gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P-PSDB; ABU38411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ds;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200277183-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   03-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Antisense;
                                                                                                                                   568
                                                                     788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACA42281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ELIT-)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ά'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wang
Wall
                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                     ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA2281

ACA
                                                       요
                                                                                                                                   ð
                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                              윤
```

```
---GGCÁÁCGCAAT 3330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3763 ACCATCAGCGGCACCGCGAGGCCGGGGCCAAGTGATCCTCACCGACGGCAACGGCAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ACCGCCGAGCCGGC 3297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3331 CCGATCGCCGAGGTCACCGCCGACGCCAGCGGCAACTGGACCTACACCCCGTCCACGCCG 3390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3703 gecaecaeggagadegeggagegecececececeggagaregarecaaecaaecaecae 3762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       155
                                                                    drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyAlaIleLeuArg---SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuPro 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116 PheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspPro 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----ProThrAspThrThrGlyTyrGlnSer 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProVal --- PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20
strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31 IleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----AsnGly---ValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATCGCCAACGGTACTGTGGTCAACGTGGCCCCAGGACGCCTCCGGTAACAGCAGCCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaAlaValGlyMetTyrThrAsnSer------TrpAspProAsnAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------AspileGinGlyValValTrpValAlaPheAspLysSerSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Leu-----GlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsn---Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TrpileProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr------
                                                                                                                                                                                                                        879 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---GATGCCGGCGGCAACCCGATAGGCAGGTCACCGCCGAC-----
                                                                                                                                                                                                                                                                             7407
205
88
314
292
50
                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Grcaaccrgagcaarggcagcagccrcagcggc----
                                                                                                                                                                                                                                                                             Length:
Matches:
                                                                                                                                                                                                                        Sequence 7407 BP; 1387 A; 2763 C; 2378 G;
                                                                                                                                                                                      ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                             (1-7407)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3298 AGCACGGTGATTCTCACCGAC-----
                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-3 (1-740) x ACA42281
                                                                                                                                                                                                                                                                             .1.36e-05
238.00
32.59%
22.80%
5.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3625 CCG------
                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              186
                                                                                                                                                                                                                                                                                                                                                        Query Match
DB:
   88888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

4813 GGCCAGGTCACCGCCGACGCCACGCGCAACTGGAGCTTC	Db 4852 ACGCCCGGCACGCCGACGCCTCGGTGATCAATGCGCTGGCCCAGGACGCCGC 4911 Qy 517GlnGlySerGluProGlyGlyValThr525	Db 4912 GGCAACAACAACCACCACCAGCGCCACCACCACCAGCGCCCAGCAG	Qy 526ThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrp 541	Qy 542 AlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer 559	560 TrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro	5089 TGGAGCTTCACGCCCGGCACGCCG	Qy 580 LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArg592	Qy 593SerThraspGlyGlyValThrPheGlnProValAlaAlaGlyLeuPro 608	Oy 609SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAsp 626	5218 ATCGACCCGAGCAACGGTGTCGAACTCAGCGGCACCGCGAACCCGGGCGTCCGGGTGATC	Qy 627 LeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrp 645	Qy 646 SerAlaileThrGlyValSerSerAlaValAsnValGlyPheGlyLys 661	5559 ROLLICACOCCOGOCACOCCOCCOCCACOCOCCOGOGOCCOCCOCCOCCACOCACCOCCOCCACOCACCACOCACCAC	Db 5398 CCGGCCGGCAATACCAGCGGCCAGCACCACGGTGGACACGGTGGCTCCGGCCACG 5457	Qy 682 GlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHis 701	Qy 702 GlnTyrGlyAsnTrpGlyGlnAla1leThrGlyAspHisAlaAsnLeuArgArgValTyr 721	Qy 722 IleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly 740	On the second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second second se	ACC53398 ID ACC59398 standard; 7407 BP. XX	AC ACC59398; XX DT 28-AUG-2003 (first entry)	XX DE Microbial resistance gene PA1874 coding sequence.	XX XX XX	OS Unidentified. XX PN W02003041483-A2.
3823 CCGATCGGCGAAACCACCGCCGACGGCAACTGGAGCTTCACGCCGGCACGCCG 3882 238	CTGGCCAACGGCACGGTGAACGCCGTGGCCCAGGACCCTGCGGGCAATACCGGCCCG	240		4003 GGCAACCTGGTCAACGGTACCGCGGGCACCGTGACCTTGACCGGCAAC 4062 269 SerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArglle 285 	4063 GGCAACCCGATCGGCCAGACCGCCGATGGCAACTGGAGCTTCACGCCCGGC 4122 286 SerProValProSerThrAspThrAlaAsnAspTvrPheGlvTvrSer 301	TCGCAACTGCCCAACGGCACCGTGATCAACGCGAGCGAGC	302. GlyLeuThrileAspArgGlnHisProAsnThrileMetVal 315		4243 AGCAACGGTTCGGTGATCAGCGGCACCGCGGACGCCGCAACACCATCATC 4296 331 ThrAspG1vG1v			355 LeuAspileSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProPro 372	AGCAATGTCGACAGTGCGCCGGTGATCACTGTGGATGGCTTGGCCCCCGGCGGCGCCCC	373 ValFroSerFroLysLeuGlyTrpMetAspGluAlaMetAlaileAspProFheAssSer 392 	393 AspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeu 409 ::: :::		4546GATGGCGGCGGCAACCCGATCGGCCAGGCCA	4585 GGCAGCGGCAACTGGACGTTCACCCCGAGCACCCGGTGGCCAACGGCACCGTGATCAAC 4644	446 SeralaLeuGlyAspLeuGlyGlyPheThrHis	460 ThrhlaValProSerThrllePheThrSerProValPheThrThrGlyThrSer 477	4/05 AICHCLCCGCCGGCGCCGGTGAICAATCCGGCGATGGAGTCGTCATCAGCGGTACG 4/61 478 VallaentvralaCluLoubenDrocertlaTlaVallarahlaGarGhaganDrocer 407		498 SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe 516
d &	do g	Š 8	કે દ	8 8	음 6	3 A	8 8	8 8	음 6	3 合	ò á	8 8	අ ර	å 8	8 8	8 &	8 &	d G	\$ B	à t	S &	3 A	ð

```
3763 ACCATCAGCGGCACCGCGGGCCGGGGCCAAGTGATCCTCACCGACGGCAACGGCAAC 3822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCGATCGGCGAAACCACCGCCGACGGCAGCGGCAACTGGAGCTTCACGCCCGGCACGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1883 CTGGCCAACGGCACGGTGGTCAACGCCGTGGCCCAGGACCCTGCGGGCAATACCGGCCCG 3942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3943 CAGGGCAGCACTACCGTGGACGCGGTGGCGCCGAACACGCCTGTGGTCAATCCGAGCAAC 4002
                                                                                                                                                                                                          HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr-----AspGlySer 268
                                                                                                                                                                                                                                                                                                                       ----GGCAGCGCCAACTGGAGCTTCACG
                                                                                                                                156 AspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAla
                                                                                                                                                                                                                                                                                  -- AspileGinGlyvalvalTrpvalAlaPheAspLysSerSerSerSer
                                                                                                                                                                   ----decacccccreecc
                                                                                                                                                                                                                                                                                                                                                              Leu-----GlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsn---Asn
                                                                                                                                                                                                                                                                                                                                                                                                                                    ProVal --- PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --PhelleProHisLysGlyValPheAspProValAsn
                                                      AsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerGlyAspValTrpLysPheSerValThr---SerGlyThrTrp.
     Asn---
                                                                                                                                                                       ပ္ပံ
                                                                                                                                                                                                            176
                                                                                                                                                                                                                                               3643
                                                                                                                                                                                                                                                                                                                                                              202
                                                                                                                                                                                                                                                                                                                                                                                                                                        219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3823
                                                                                                                                                                                                                                                                                    186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                      ò
                                                                                       용
                                                                                                                                ò
                                                                                                                                                                 a
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                            용
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                 capable of altering the sensitivity of a microorganism to an antimicrobial agent by employing efflux pumps comprising polypeptides encoded by the following genes: PA1874, PA4142, PA2389, PA1876, PA4143, PA2390 or PA163. The method is useful for identifying modulators of microbial resistance of an organism in a biofilm. The methods are also useful for identifying genes that encode proteins that play a role in biofilm resistance. The method is particularly useful for screening compounds or discovering compositions that will inhibit biofilm formation and overcome their resistance mechanisms. These methods are particularly useful in medical, industrial or natural settings, where formation of biofilms can have serious negative consequences and result in high costs both in human health and economic terms. The present sequence is a
                                                                                                                                                                                                                                                               Identifying modulators of microbial resistance of organisms in biofilms, e.g. inhibitor of biofilm formation, by employing expression controls, or efflux pumps containing polypeptides, of genes associated with biofilm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   coding/control sequence shown in the exemplification of the invention
                                                                                                                                                                                                                                                                                                                                                                                                 present invention relates to a method of identifying a compound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 7407 BP; 1387 A; 2763 C; 2378 G; 879 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                              Disclosure; Fig 7; 102pp; English.
                                                      18-SEP-2002; 2002WO-US029565
                                                                                           18-SEP-2001; 2001US-0323241P
                                                                                                                                (DART-) DARTMOUTH COLLEGE
                                                                                                                                                                                                          WPI; 2003-468567/44.
P-PSDB; ABP59933.
                                                                                                                                                                   Mah T;
                                                                                                                                                                     g,
                                                                                                                                                                                                                                                                                                                         resistance
                                                                                                                                                                     O'toole
```

3331 CCGATCGCCGAGGTCACCGCCGACGCCAGCGGCAACTGGACCTACACCCGGTCCACGCCG 3390 3451 CCGGCGACGGTGACCGTCGATTCCAGCGCCGCCGGCGCCGGTGATCAACCCGAGCAAC 3510 116 PheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspPro 135 82 3391 ATCGCCAACGGTACTGTGGTCAACGTGGCCCCAGGACGCCTCCGGTAACAGCAGCCCG ValAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly ---ACCGCCGAGCCGGGC 31 IleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArg ----TrpAspProAsnAsp -----AsnGly---ValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrp GlyAlaIleLeuArg---SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuPro TrplleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr-------7407 205 88 314 292 50 Length:
Matches:
Conservative:
Mismatches:
Indels: AlaAlaValGlyMetTyrThrAsnSer------|||| ::: |||| GTCAACCTGAGCAATGGCAGCAGCCTCAGCGGC Gaps: x ACC59398 (1-7407) ::: 3298 AGCACGGTGATTCTCACCGAC-1.36e-05 238.00 32.59% 22.80% 5.90% US-09-917-376-3 (1-740) Similarity: Percent Similarity: Alignment Scores: Pred. No.: 7 21 83 97 99 Query Match: DB: Best Local ò ò 셤 ઠે 셤 ò 셤 ઠે 셤 ò 유

GCCAACCCGATCGGCCGACCCCGATGGCACCGCCAACTGGAGCTTCACGCCCGGC 4122

1063 286 4123

셤 ઠ 요 ઠે a Š

269

251

rciccaacreciccaaceecaccerearcaacereacedeceaeceaceeceeceecearace 4182

SerProValProSer-----ThrAspThrAlaAsnAspTyrPheGlyTyrSer

301

1183 AGCCTTCCCGCTACCACGACGGTGGATTCCTCGCTGCCGTCGATCCCGCAGGTGGATCCG 4242

302 GlyLeu-----ThrIleAspArgGlnHisProAsnThrIleMetVal---

-----AlaThrGlnIleSerTrpTrpProAsp-----ThrIleIlePheArgSer 330

Agcaacegttcegtgárcáccecaccececaccecacacacáccarcárcatc----

4243 331 1297 343 4357 355 4417 373 4477

유

ò a ò 셤 ò 셤 Š 셤

316

ACCGATGGCAACGGCAACCCGATTGGCCCAGGTCACCGCCGACGGCAGCGGCAACTGGTCC 4356

ThraspGlyGly-------ThraspGlyGly----------AlaThrTrpThrArglleTrpAsp

--SerTyrProAsnArgSerLeuArgTyrVal

TrpThr--

LeuAspIleSerAlaGluProTrpLeu-----ThrPheGlyValGlnProAsnProPro

ValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaileAspProPheAsnSer 392

------ATCGATCCGAGCAACGC 4497

393 AspArgMetLeuTyrGlyThr-----GlyAlaThrLeuTyrAlaThrAsnAspLeu 409 ::: |||||||

```
5088
                                                   4761
                                                                                                                                                                                                                                                                                                                      ------GCGCBAGCCGGGCCACGGTGATCCTCACCGACGCAACGGCAACCGCATC 4812
                                                                                                                                                                                                                                                                                                                                                                                       4851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----- CCGGTGATCAATCCCAGCAAC--- 5478
                                                                                                                                                                                         1645 GCCGTGGCCCAGGACCCGGCCGGCAATACCAGCGTCCGGCCAGCGTCACCGTCGATGCC 4704
                                                                                                                                                                                                                            477
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1852 ACGCCCGGCACGCCGCTGGCCAACGGCTCGGTGATCAATGCGCTGGCCCAGGACGCCGCC 4911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1912 GGCAACAACÁGCAGTCCCACCAGCGCCÁCCGTCGACTCGCTGGCGCCAGCAGCCCCGGTG 4971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                608
                                                                                                                                                                                                                                                                                                                                                         516
                                                                                                                                                                                                                                                                                                                                                                                                                       516
                                                                                                                                                                                                                                                                                          497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHis 701
                              410 ThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGlu
                                                                                               ------ProSerGlyAlaProLeuile
                                                                                                                 SerAlaLeuGlyAspLeuGlyGlyPheThrHis-------AlaAspVal
                                                                                                                                                                                                                       460 ThralaValProSerThrIlePhe-----ThrSerProValPheThrThrGlyThrSer
                                                                                                                                                                                                                                                       1705 ATCGCCCCCCCCCCCCCCCCCCCCATCAATCCCAATGGAGTCGTCATCACCGGTACG---
                                                                                                                                                                                                                                                                                        478 ValAspTyrAlaGluLeuAsnProSerIlelleValArgAlaGlySerPheAspProSer
                                                                                                                                                                                                                                                                                                                                                                                       -----GTCACCGCCGACGCCACCGCAACTGGAGCTTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------ThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4972 ATCGATCCGAGCAACGGTAGCGTGATCGCCGGTACCGCCGAGGCTGGTGCCACGGTGATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  542 AlaProGlyAspProGlyGlnProVal -----ValTyrAlaValGlyPheGlyAsnSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CTCACCGACGCCAACGCAACCCGATCGGCCAGGTCACCCCCGATGGCAGCGGCAAC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------crércaarégeaceregreangegeeceaegacereec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---SerThrAspGlyGlyVal-----ThrPheGlnProValAlaAlaGlyLeuPro---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5158 GGCAACAGCGGCCCGGTCAGCACCACGGTGGACGCGGTGGCCCCGGCCACCCCGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5218 ATCGACCCGAGCAACGGTGTCGAACTCAGCGGCACCGCCGAACCCGGCGCGTCAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSer---ThrAsnGlyGlySerSerTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerAlaIleThrGlyVal------SerSerAlaValAsnValGlyPheGlyLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AGCTTCACGCCGGGCGCCGCTGGCCAACGGCACGCTGGTCAATGCCCGTGGCCCAGGAC
                                                                                                                                                                                                                                                                                                                                                      SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          crcaccaargecaargecaarccgarcggccagaccrrgccgacgcagcggcaacrgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3398 CCGGCCGGCAATACCAGCGGCCCGGCCAGCACCACGTGGACACGGTGGCTCCGGCCACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThr
ACCGAG---ATAAGCGGTACCGCGGAGGCCCGCGCGACGGTGATCCTCACC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               580 LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArg-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---GlnGlySerGluProGlyGlyValThr-
                                                                                           430 ThrAlaValAsnAspLeuIleSerPro-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                recaeccecececece
                                                                                                                                                                                                                                                                                                                                                                                      1813 GGCCAG-----
                                                                                                                                                                                                                                                                                                                                                        498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5338
                                                             1546
                                                                                                                                                           446
                                                                                                                                                                                                                                                                                                                      1762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5458
                                                                                                                                                                                                                                                                                                                                                                                                                     516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           662
                                                                                                                   셤
                                                                                                                                                                                 g
                                                                                                                                                                                                                                                  g &
                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                            8 %
                                                                                                                                                                                                                                                                                                                                                                                                                    8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                            8
                                                   셤
                                                                                 &
                                                                                                                                                         ઠે
                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
the foll antiennes sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated containing the vector; (3) an isolated are:

(2) proliferation or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular continuation or the activity of a gene in an operon required for proliferation or the activity of a gene in an operon required for the proliferation, (7) identifying a compound that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound; sactivity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent of product is overexpressed or underexpressed; (12) determining the extent compound; sactivity; (11) a culture comprising strains in which each of the strains is present in a culture or collection of the strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids are useful for identifying proteins or screening for homologous nucleic acids are useful for identifying the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary and the careary
                                     GlnTyrGlyAsnTrpGlyGlnAlaileThrGlyAspHisAlaAsnLeuArgArgValTyr 721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               or screening
                                                                                                          one of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zyskind JW;
Xu HH;
                                                                                 IleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New antisense nucleic acids, useful for identifying proteins or s for homologous nucleic acids required for cellular proliferation isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        invention relates to an isolated nucleic acid comprising any
                                                                                                                                                                                                                                                                                                                                                                                   ds; prokaryotic essential gene; cell proliferation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ohlsen KL,
Forsyth RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14; SEQ ID NO 11269; 1766pp; English
                                                                                                                                                                                                                                                                                                                                             Prokaryotic essential gene #5056.
                                                                                                                                                                                                                 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malone C,
Carr GJ,
                                                                                                                                                                                                               ACA23399 standard; DNA; 7305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-MAR-2002; 2002WO-US009107.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-FEB-2002; 2002US-00072851.
                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (ELIT-) ELITRA PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2003-029926/02.
P-PSDB; ABU19529.
                                                                                                                                                                                                                                                                                                                                                                                                       design; gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Borrelia cepacia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200277183-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAR-2001;
                                                                                                                                                                                                                                                                                                   19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     03-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                     Antisense;
702
                                                                                                                           5521
                                         5479
                                                                                   722
                                                                                                                                                                                                                                                         ACA23399;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'n'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim
                                                                                                                                                                                                                                                                                                                                                                                                         drug
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wang
Wall
                                                                                                                                                                     RESULT 10
                                                                                                                                                                                           ACA23399
                                     a
                                                                                                                         셤
                                                                                                                                                                                                               ਨੇ
                                                                                 ò
```

```
TCGGCCACGGCGGAAGCGGGGGGGGGGGCGACGATACGAAAGGCGACGGAAGGCCG 1053
for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is one of the trarget prokaryotic essential genes. Note: The sequence data for this patent dinct form part of the printed specification, but was obtained in electronic format directly from WIPO at the wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGly----- 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsn-----PhePro 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSer 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   224 ArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHis 243
                                                                                                                                                                                                                                                                                                                                                                          576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----AspGlyAlalleLeu-----ArgSerSerAspGln 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GlyAsnMetProGlyArgGlyMetGlyGluArgLeu 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeu 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               817 TCCGGGAAGTGGACATTTACGCCGACCACG---GGTTTGGGCGAGGGCGCGCACTCGATT 873
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    95
                                                                                                                                                                                                                                                                                                                                                                                                       LeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAla 73
                                                                                                                                                                                                                                                                                                                                          -----IlePro 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1171 TCTGCTTCGGTAACGGTGACGGCGATACGACGCCACCGGGTGCGCCCGGTGATC----
                                                                                                                                                                                                                                                                                                                                                                          GCGAACGGCCGCTGGTCGTACACGCCCGGTACGTTGCCGGAGGGCGCCCCATCAGATCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                     GTGACGCAGACCGACAGTAGCGCGAGCGAGCTGAGCACGGTGGGGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GATACGATCGTACAGAAAACGCCTGCGATTACGGCCGTCAGCGACGATGCACAACATCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 757 GGTAGCAGGGTGAGCGTGTATGACGGCACGACGCTGCTGGGAAACGACGACAGCGGGATTCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAsp----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---AlaValGlyMetTyrThrAsnSerTrpAspProAsn-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----ileGlnGlyValValTrpVal---AlaPheAspLysSerSerSerLeuGly
                                                                                                                                               Sequence 7305 BP; 1395 A; 2299 C; 2566 G; 1045 T; 0 U; 0 Other;
                                                                                                                                                                                             7305
185
72
303
243
                                                                                                                                                                                                                            Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AspProlle-----AsnThrAsnLysValTrpAla
                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                           US-09-917-376-3 (1-740) x ACA23399 (1-7305)
                                                                                                                                                                                           0.00414
199.50
32.00%
23.04%
4.94%
                                                                                                                                                                                                                                                                                                                                          AlaAsnGlyArgTrp
                                                                                                                                                                                                                                           Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                            Percent Similarity:
                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1054
                                                                                                                                                                                                                                                                                                                                          47
                                                                                                                                                                                                                                                                                                                                                                                                       54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1225
                                                                                                                                                                                                                                                                                                                                                                                                                                       577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
   8888888888888
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

6 6.	244 LysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGly 263
ò	264 ProSerGlySerSerGly 270
qa	1330 ACGTCAGCGTGTATGACGGCACGACGACGTTGTTGGGTACGACGACGACGACCGAC
& &	271 AspValTrpLysPheSerValThrSerGlyThrTrpThrArgileSerProVal 288 :::
ઠે	89 ProSerThrAgpThrAlaAsnAspTvrPheGlyTyrSerGlyLeuThrIle 30
q	
ζ	306 AspArgGlnHisProAsmThrIleMetValAlaThrGlnIleSerTrpTrpProAsp 324
DP	1507 GATACGACAGCTCCGTCCGACAGTGAACGCGACCGACGGT 1551
ò	325 ThrilellePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThr 344
qq	1552 ACGTCGCTGTCGGGCACGGCGGAGCGGCACCATCAACATCGATACGAACGGCGAC 1611
ò	5 SerTyrProAgnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThr 364
අු	-
È	oLysLeuG
qq	1654 TGGACGTACACGCCCTCGACCCCGCTGCCGGCT
È	385 MetalalleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyr 404
qq	1687GGCACGGTGATTGGCGTGACG 1707
ò	4
qq	
δ	
Dp	i
ò	445 IleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSer 464
QQ	1789 ATCGGTACGGTGCGGACGATGCGGCCTTGTTGTTGGTGCGATTGCCTCTGGCGGCAGT 1848
٥٨	rThrGlyThrSerVa.
qq	
ò	485 ProSerIleIleValArgAlaGlySerPheAspProSer 497
qq	1894 AGCACGGTGAGCGTGTATGACGGCACGACGCTTGCTCGGCACGACGACGGCGGACCGACC
ò	oAsnAs
Ωp	
à	517Thr 526
Db	1969 ACGCCGACTACTGCACTGGGTGAGGTGCACACAGCCTGACCGTCACCGCAACGGATACG 2028
ò	527 GlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspPro 546
QQ	2029 GCCGGCAACGTCAGTTTCCGAGCACGGCCTTCGATCTGACGATCGAT
ò	547 GlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGly 565
g	2089 gcgarrcccacadrcaaccccacccaccccarccarccaccacccccccc
ò	566 ValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeu 585

 	SerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAla	CCCAGCGCGCGTGG	606 GlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGly 625 :: :: 2242 CCGCTGCCGGCTGGTACGGTGTGACGGCGTGACGATGCGGTGGGGAGCAAC 2295	626 AspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrp 645	2295 2295	646 SerAlalleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGly 665	2296ACGGGCCCGTCTGCCTCGGTAACGGTGACGGTGACACGACGCACGGGGT 2346	SerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThr	GCGCCGGTCATCGGCACCGTGACGATGACGTGCTTCTTTTT.	osz otyanitananyinegsetabbabbysetyintintipvallentieabn 697 	698 AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnLeu 717	2443	ArgArg	2473GTGTATGACGCCACGACGTTGTTGGGTACGACGGCCGAC 2514	738 ProserGly 740 2515 CCGTCCGGA 2523	RESULT 11 ACA24001/c ID ACA24001 standard; DNA; 5295 BP.	ACA24001;	19-JUN-2003 (first entry)	Prokaryotic essential gene #5658.	Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.		W0200277183-A2.	03-0CT-2002.	21-MAR-2002; 2002WO-US009107.	21-MAR-2001; 2001US-00815242. 06-SEP-2001; 2001US-00948993. 25-07-2001; 2001US-03423P. 08-FCB-2002; 2002US-0072851. 06-MAR-2002; 2002US-0362699P.	(BLIT-) ELITRA PHARM INC.	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	WPI; 2003-029926/02. P-PSDB; ABU20131.
QQ	ď	Q Q	රු පු	à	QQ	δ	q	à	2 2	g G	δ	q	ò	g	\$ a	RE: AC! ID	ΧX	ΧĘ	X 23 :	X & &	XX SO	X M	X E ;	FP.	4 K K K K K K	X E S	Y II II X	888

```
The invention relates to an isolated nucleic acid compileration of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
colypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) a natibody capable of specifically binding
the polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) a natibody capable of specifically binding
the polypeptide or its fragment whose expression is inhibited by the
cartisty of a gene in an operon required for
proliferation or the activity of a gene in an operon required for
proliferation, or that has an activity against a biological pathway
compound or the activity of a gene or the biological pathway
in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (2) determining the extent
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (2) determining the extent
compound's activity; (12) a culture comprising strains in which the gene
compound's activity; (13) a culture comprising strains in the compound that inhibits the
compound's activity; (12) a culture compressed; (2) determining or an organism. The antisense nucleic acids are useful for proliferation of an organism. The antisense nucleic acids are useful
control discovery programs, or for screening homologous nucleic acids
crequired for proliferation i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98 AlaileLeu-----ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeu 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78 ThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle
                                                                                                                                                                         The invention relates to an isolated nucleic acid comprising any
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 5295 BP; 760 A; 1863 C; 1647 G; 1025 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115 ProPheLysLeuGly------
                                                                                                                        Claim 14; SEQ ID NO 11871; 1766pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-3 (1-740) x ACA24001 (1-5295)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0.0056
195.00
33.11%
23.43%
4.83%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
```

쇰

ò

ઠે

141 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr

161 TrpSerGlnMetThrAsn-----PheProAspValGlyThrTyrIleAlaAsnPro 177

= :: =

8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6

Second	ACA40315 ACA40315 ACA40315 ACA ACA40315; XX AC ACA40315; XX AC ACA40315; XX AT 19-JUN-2003 (first entry) XX XX XX XX XX XX XX XX XX XX XX XX XX
888888888888888888888888888888888888888	
GTCAACATCGATACGAAGGCGACGGTACGCGGACGCCACCGTGACGGCTGACCCC ThraspThrThrGlyTytGlnSerAsp	4178GGCACGGTGATTGGCGTGACGGCTACC

8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6

```
the 6113 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid;

(2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation or the activity of a gene in an operon required for proliferation, (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway cequived for that the set compound that inhibits proliferation; (8) identifying a gene required for required gene or its gene product lies or agene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture complising strains in which the gene product is oversxpressed; (12) determining the extent the compound's activity; (11) a culture compound; (12) determining the extent the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the compound of the 
                                                                                                                                                                                                                                                                                                                                                                                                                                                   screening
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          strains, or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    invention relates to an isolated nucleic acid comprising any one of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the
                                                                                                                                                                                                                                                               Zyskind JW;
Xu HH;
                                                                                                                                                                                                                                                                                                                                                                                                                                New antisense nucleic acids, useful for identifying proteins or scre
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                               Ohlsen KL,
Forsyth RA,
                                                                                                                                                                                                                                                            Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 14; SEQ ID NO 28185; 1766pp; English
                                                                                                                                                                                                                                                            Malone C,
Carr GJ,
                                                      06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342922P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
                           21-MAR-2001; 2001US-00815242
                                                                                                                                                                                                     ELITRA PHARM INC.
                                                                                                                                                                                                                                                            Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                   P-PSDB; ABU36445
                                                                                                                                                                                                                                                               Wang L,
Wall D,
                                                                                                                                                                                                     (ELIT-)
```

required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in Sequence 9903 BP; 2129 A; 3380 C; 2727 G; 1667 T; 0 U; 0 Other; 9903 194 94 323 213 Matches: Conservative: Mismatches: Indels: Length: ftp.wipo.int/pub/published_pct_sequences directly from WIPO at 0.018 192.00 35.04% 23.60% 4.76% Alignment Scores: ectroni

drug discovery programs, or for screening homologous nucleic acids

(1-9903) US-09-917-376-3 (1-740)

Percent Similarity: Best Local Similarity:

Query Match:

Gaps:

	27	7449	41	7509
02-03-311-310-3 (I-740) X ACA40313 (I-9903)	10 AsnvalAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGly 27	7399 AACGCGAACATCGGCGCGAAACATCGGCGACTTTAACGTCGGGATCGCA 7449	28AlaProGly1leLeuTyrValArgThrAspIleGlyGlyMet 41	7450 AACACCGGTCCGGGGCTAACGGCGGCTGTCAACAACATCGGTATCGGCAACACCGGCAAC 7509
		•		•
200	ò	g	ò	Ωp

ò	42	TyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp 60
qq	7510	
ò	61	AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsp74
Db	7555	Ü
à	75	ProlleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp 93
ор	7615	GGCCCGCTGAACGCCGGCATCGCCAACATGGCCTGTTCAACCTGGGCGAC 7665
ò	94	Alaile
qq	7666	AACAACTITGGC7677
ŏ	114	LeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal 133
d d	7678	
à	134	AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrp 152
QQ	7714	AACACCGGCAACAACAACAACGGCTTGTTCAACACCGGCAACAACAACGTCGGCATCTGG 7773
à	153	ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr 172
Dp	7774	CTGACCGGCGACGGCTTGTCCGGCTTCAGCTCCCTGAACTCCGGCGCGCGC
ò	173	TyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrp 192
Db	7834	TTCTTCAACTCCGGCACCGCCAACACCGGC7863
à	193	ValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThr11ePheVal 211
QQ	7864	TTGTTCAACTCCGGCAACGGCAACGGGCTTGTTCAACTCGGGCACGGCAACGTC 7920
à	212	GlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGln 231
Op	7921	GOCATCOCCAACATGGGCACCGGCGGCTTCGGCGTCGGCCTATCCGGCGACAGGCCAGGTG 7980
λ̈́o	232	AlaValProGlyAlaProThrGlyPhelleProHisLySGlyValPheAspProValAsn 251
Db	7981	GGCATCGGCGCACCAACTCGGCAGTTTCAACATCGGCTTGTTTAACTCGGGCACC 8037
à	252	HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGly-As 271
Db	8038	GGCAATGTCGGCATCGGCAACTCGGCCACGCCCATCGGCAACGCGGCACCGGCACC
à	271	pvalTrpLysPheServalThrSerGlyThr
QQ	8098	GGCAACACCGGCATCGGAAAAACGGCAACTACAACGGGCTTGCTCAACGCGGGCCTG 8157
à	282	pThrArglleSerProValProSerThrAspThrAlaAsnAs 296
qa	8158	Grcaacaccegearceceaaccegaaccacaacaccegecererreaacarcegeacc 8217
š	296	pTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAl 316
Db	8218	TTCAACACCGGCATCGCCAACCCGGGCCACTACAACACCGGCTCCTACAACACCGGTAGC 8277
à	316	aThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThr 331
qq	8278	TACAACACGGGGATGGGAAACGCCGGGAGCTACGGCGCGCGC
ò	332	
qq	8338	ATGAACAACGGCTTGCTCTGGCGCCCGACCGGCAGGGCCTGCTGGCGGCCAACTACACC 8397
ò	346	TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGl 360
QQ	8398	ATCACCATCGAGGGACCTGCCGCGTTCCTCAATGTCGACATCCCGGTCAACATCCCCATC 8457
ð,	360	uProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTr 380

	346	1 3	360
•	26280	ATCACCATGGAGGGACCTGCCGCGTTCCTCAATGTCGACATCCCGGTCAACATCCCCATC 20	26221
	360	<pre>trpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTr 3 </pre>	380
_	26220	7	26170
	0	etAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyTh 3	399
_	26169		26114
	399	rGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGln11 4: 	118
	418	1 4	3 8
•	53	CTTCGGCCCCTCGCCGGCGATCAACCTCAACATCGGCAA 2	26012
	438	OProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAs 4!	158
_	26011	2	25969
	458	rIlePheThrSerProValPheThrThrGlyThrSerVa 4	78
	25968	7	25952
	478	IleIleValArgAlaGlySerPheAspProSerSe 4	98
_	25951	AGCGCCCGGCTTCTTCAACGCCACCAC	25904
	498	snAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGl	18
•	25903	CGGCCCGTCGTCGGGCTTCCTCAACTGGGGTGCTGG 2:	25868
	518	roGlyGlyValThrThrGlyGlyThrValAlaAl	532
_	25867	rcgggccrcracaacrrcgccacrag 2	25808
	532	aSerAlaAspGlySerArgPheValTrpAlaProGl 50	544
_	25807	2	25753
	544	ы Б	960
	25752	arcregecarera 2	25724
	260	GlnIleArgSerAspArgValAsnProLy 5	280
_	25723		25673
	280	SThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVa 59	298
_	25672	2 2	25639
	598		617
_	25638	(i)	25583
	618		630
_	25582		25523
		aAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerA	647
_	25522	7	25463
	647	laileThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerS 66	293
_	25462	~	25414

```
The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibites proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated opypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; (7) identifyly of a gene in an operon required for proliferation; (7) identifyly a activity against a biological pathway required for proliferation, or that has an activity against a biological pathway and agene required for cellular proliferation; (8) identifying a gene required for cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an
                                                                                                                                                                        25361 TCGGCATCGGCAATACCGGCACCGGCAACATCGGCTTCGGCAACACCGGAAACGGAAACA 25302
                                                                                                                                                   rpGlyGlnAlaIleThrGlyAsp--HisAlaAsnLeuArgArgValTyrIle-GlyThrA 725
erTyrProAlaValPheValValGlyThrileGly---GlyValThrGlyAlaTyrArgS 686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          or screening
                                                                                                                                                                                                                                                686 erAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnT
                                                                                                                                                                                                                         snGlyArgGlyIleVal-----TyrGlyAspIleGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zyskind JW;
Xu HH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ds; prokaryotic essential gene; cell proliferation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        antisense nucleic acids, useful for identifying proteins homologous nucleic acids required for cellular proliferat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ohlsen KL,
Forsyth RA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Haselbeck R,
Yamamoto R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 14; SEQ ID NO 26288; 1766pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Prokaryotic essential gene #20075
                                                                                                                                                                                                                                                                                                                                        ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malone C,
Carr GJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-0362851.
06-MAR-2002; 2002US-0362699P.
                                                                                                                                                                                                                                                                                                                                        ACA38418 standard; DNA; 3324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-MAR-2002; 2002WO-US009107
                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (ELIT-) ELITRA PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zamudio C,
Trawick JD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mycobacterium bovis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2003-029926/02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            drug design; gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; ABU34548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200277183-A2
                                                                                                                                                                                                                                                                                                                                                                                                               19-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   03-OCT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Antisense;
                                                                                                                                                                                                                           725
                                                                                                                                                                                                                                                                                                                                                                          ACA38418;
                                                                                                                                                   904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wang L,
Wall D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New
                                                                                                                                                                                                                                                                                                                   ACA38418
                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                        g
                                  g
                                                                          ð
                                                                                                       ద
                                                                                                                                                   ò
```

```
organism acts; (9) manufacturing an antibiotic; (10) profilling a compound's activity; (11) a culture comprising strains in which the gene product is overspressed or underexpressed; (12) determining the extent ct owhich each of the strains is present in a culture or collection of the strains is present in a culture or collection of corrections of an organism. The antisense nucleic acids are useful for intentifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational carp programs, or for screening homologous nucleic acids required corrections or acreening for homologous nucleic acids required crequired for proliferation in cells other than S. aureus, S. typhimurium, C. prokunctic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in celectronic format directly from WiPo at.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              111 IleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArg 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AACGCGAACATCGGCGCGAAACATCGGCGAC-----TTTAACGTCGGGATCGCA 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TÄCAACATCGGTGTCGGCAÄCACCGGTAACTACAAC--------ATCGGCTTC 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProAsnAspGly---AlaIleLeuArgSerSerAspGlnGlyAla-----ThrTrpGln 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Leu-----TrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhePro 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AspvalGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGln 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -------Trefrcaacrecescacescacescescificarreces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ThrilePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGly 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232 AACACCGGTCCGGGGGTAACGGCGGCTGTCAACAACATCGGTATCGGCAACACCGGAAC 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGCCCGCTGAACGCCGGC----ATCGCCAACATGGGCCTGTTC---AACCTGGGCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----AlaProGlyIleLeuTyrValArgThrAspIleGly------GlyMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TyrArgTrpAspAlaAlaAsn---GlyArgTrpIleProteuLeuAspTrpValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---ProlleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----ACCGGCGACGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln---AlaSerLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnValAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGly-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AACAACATCGGCTTGTTCAACACGGCAACAACAACGTCGGCATCGGGCTG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LeuAlaValAspProAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 AsnAsnTrpGlyTyrAsnGlyVal---ValSerIleAlaAlaAsp-------
                                                                                                                                                                                                                                                                                                      Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AACTCCGGCACC----GCCAAC------ACCGGC---
                                                                                                                                                                                                                                                                                                    0 12 0
                                                                                                                                                                                                                                                                                                                                                           3324
205
89
313
221
                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                    .,
H
                                                                                                                                                                                                                                                                                                      BP; 726 A; 1184 C; 886 G; 528
                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-3 (1-740) x ACA38418 (1-3324)
                                                                                                                                                                                                                                                                                                                                                         0.0169
184.00
35.59%
24.82%
4.56%
                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                      Sequence 3324
                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           559
                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                              ..
No.:
           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

169 AspValGlyThrTyrIleAlaAsnPrOThrAspThrThrGlyTyrGlnSerAspIleGln 188 189	292 pThraladsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnTh	
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 6 8 6 1	8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6
WP AA199683 18 1800001 1910000 WP AA199683 20 2000001 2010000 WP AA199683 21 2000001 2010000 WP AA199683 22 2000001 2010000 WP AA199683 23 2000001 2010000 WP AA199683 24 2400001 2010000 WP AA199683 25 2500001 2010000 WP AA199683 26 2600001 2010000 WP AA199683 28 2800001 2010000 WP AA199683 29 2900001 2010000 WP AA199683 30 2000001 2010000 WP AA199683 31 3100001 3110000 WP AA199683 31 3100001 3110000 WP AA199683 34 3400001 3110000 WP AA199683 34 3400001 3110000 WP AA199683 34 3400001 3110000 WP AA199683 34 3400001 3110000 WP AA199683 34 3400001 3110000 WP AA199683 34 3400001 4110000 WP AA199683 34 3400001 4110000 WP AA199683 40 4000001 4110000 WP AA199683 41 4400001 4310000 WP AA199683 42 4200001 4310000	Alignment Scores: Pred. No.: Score: Score: 184.00 Marches: 205 205 205 205 205 205 205 20	0 0 0 0 0 0 0 0 0 0

Search completed: May 12, 2004, 05:39:13 Job time: 1279.48 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on: May 12, 2004, 04:33:52; Search time 6950.51 Seconds
(without alignments)
3179.336 Million cell updates/sec
Title: US-09-917-376-3
Sequence: 1ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
Scoring table: Xgappp 10.0, Xgapext 0.5
Xgappp 10.0, Xgapext 0.5
Ygappp 10.0, Ygapext 0.5
Fgappp 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match O%
Maximum March 100%

Total number of hits satisfying chosen parameters:

55026578

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09917376/runat_11052004_114535_28249/app_query.fasta_1.1429
-DB=EST -QFWT=fastap -SUFFTX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR=10550917376 @CGN 1 1 5334 @runat 11052004 114535 28249 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THRABAS=1 -XGAPEXT=0.5 -FCAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

BST:*

1: em_estba:*
2: em_esthun:*
3: em_estin:*
4: em_estro:*
5: em_estro:*
6: em_estro:*
7: em_estro:*
10: gb_est2:*
112: gb_est2:*
113: gb_est2:*
114: gb_est2:*
115: em_estcon:*
15: em_estcon:*
16: em_estcon:*
17: em_gss_inv:*
18: em_gss_inv:*
19: em_gss_inv:*
22: em_gss_inv:*
23: em_gss_inv:*
24: em_gss_inv:*
25: em_gss_inv:*
26: em_gss_inv:*
27: em_gss_inv:*
28: gss_inv:*
28: em_gss_inv:*
29: em_gss_inv:*
29: em_gss_inv:*
21: em_gss_inv:*
22: em_gss_inv:*
23: em_gss_inv:*
24: em_gss_inv:*
25: em_gss_inv:*
26: em_gss_inv:*
27: em_gss_inv:*
28: gss_inv:*
28: gss_inv:*
28: gss_inv:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

•	Description	872104 tri	tric02	tri	tric0	tri	tric08		tric01	tri	_			BU639045 mgcw011xD		tric03	tric(tric(tric(RIDR	Mus n	_	mgurk		SP105	Mus	Mus m	AY402312 Homo sapi	Mus m	CH26	Mus		7742 Ношо	3469 Mus m	9959 Homo	07967 Ното	2937 Mus mus	1758 EST632	4011 9013896	30673 Mus mus	04484 Homo sa	18194 Homo sa	39264 Homo sa	32858 Mus mus	95 Ното ва	((((
COLUMNIA	ΠD	37210	90233	37691	90538	38071	90762	38206	36888	9898	36638	335	566	BU639045	581	176	177	343	592	191	194	AY416870	7.7	294	AW064456	84	7	AY402312		CC203127		80	7	9	35	96/	293	175	101	190	<₩	Y41819	926	K03285	229	C
	B	14	14	14	14	14	14	14	14	14	14	14	10	13	14	14	14	14	14	14	11	59	13	11	0	11	11	59	11	28	11	11	29	11	11	5	59	13	14	11	29	29	11	11	11	-
•	Length	704	761	747	814	707	782	693	751	803	738	794	546	389	713	782	719	775	929	196	3057	4355	296	4631	33	62	75	4101	65	17	78	36	0	88	28	56	67	42	2545	56	3264	61	3762	13	5	7,
Query	Match	'n	S.	4.	4.			٠				•	٠	7.6	•	7.1	٠	٠			٠		4.2	4.2	٠		٠	٠	٠	٠		•	٠		•	•	•			•	3.3	•	•	•	•	
	Score	15	15	87	87	15	15	85	361.5	61	338	338	7	306.5	8	8	82	82	78	222	183	180	171.5	170	99	153.5	23	153.5	151	148.5	-	139.5	Н	137.5	13	136.5	3	3	•	3	132	3	•	31.	31.	•
Result	No.	н	~	m	4	S	v	7	8	đ	. 10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	c 27	28	c 29	30	c 31	m	33	34	35	36	37	38	39	40	41	42	43	C 44	4

	CF872104 / 04 bp mkNA linear ES1 31-OC1-Z003	n 6 October 2003	equence.						Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;	pocrea.	
77	Tine	version	mRNA E				reesei)		ina; Sc	ae; Hy	
	MKNA	culture,	c029xo13,				choderma 1		zizomycoti	Hypocreace	
<u>;</u>	do .	ellal	le tri				i: Tri		a; Pe	les;	
Ċ	0,	reeseı myc	Hypocrea jecorina cDNA clone tric029xo13, mRNA sequence.		126786		Hypocrea jecorina (anamorph: Trichoderma reesei)		Ascomycot	Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea	₹
	,	o13.bl T.	jecorina		CF872104.1 GI:38126786		jecorina	Hypocrea jecorina	a; Fungi;	mycetidae	1 (bases 1 to 704)
1	CF872104	tric029x	Hypocrea	CF872104	CF872104	EST.	Hypocrea	Hypocrea	Eukaryot	Hypocreo	1 (раве
RESULT 1 CF872104	rocas	DEFINITION		ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM			REFERENCE

ALIGNMENTS

~

```
568 TCCAAGGTCTCGTCGTTCACGGCAACTGGGACGTACATCCCAGACCCGAGTGATTCCAAC 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hypocrea jecorina (anamorph: Trichoderma reesei)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                761
1118
37
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CA 94304,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.761
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric029xol3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genencor Intl.
925 Page Mill Road, Palo Alto,
721: (650) 846-7635
Fax: (650) 621-7817
Email: Pforemandgenencor.com
Seq primer: LT-F1 primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Palo Alto,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-3 (1-740) x CB902332 (1-761)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: Pamela K. Foreman
                                                                                                                                                                                                                                                      CB902332.1 GI:30116990
                                                                                             202 LeuGlyGlnAlaSer 206
                                                                                                                           687 ACCGGNGGAGCCACG 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2.73e-43
615.50
75.61%
57.56%
15.25%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12788920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                           RESULT 2
CB902332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                            셤
                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                ઠે
                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142 TyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrp 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98 TITICATGGAAGAAGGTCAAGCTCGGCGGCGGCGGCGGCTTCGTCCCGGCATCATCTTC 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           388 régréadacececécécaacerégrecritéaacritécecritéaagregégeraac 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 MetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               508 TACTTIGGIGCTCGCTCAGGAAACGGCCTCTGGAAGTCTACGGACGGCGCGTGACCTTT 567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThr 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                   /dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44 TrpAspAlaAlaAsnGlyArgTrpIlePro-----LeuLeuAspTrpValGlyTrpAsn 61
                                                                                                                                                                                                                                                                                                                                                                               /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 TyrThrTrpSerAsnValAlaile---GlyGlyGlyGlyPheValAspGlyIleValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  in a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 AsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProlleAsnThrAsnLysVal
                  and
                                     Analysis of the protein processing and secretion pathways
Trichoderma ressel EST dataset
Unpublished (2003)
Contact: Ralph A. Dean
Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D., Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             704
1118
37
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                          Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
                                                                                                                                                                                                                                     1. .704
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                    /db_xref="taxon:51453"
/clone="tric029xo13"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-3 (1-740) x CF872104 (1-704)
                                                                                                                                                       Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph deamoncsu.edu
Seq primer: LT-Fl primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2.4e-43
615.50
75.61%
57.56%
15.25%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment
AUTHORS
                                                                           JOURNAL
                                                                                                                                                                                                                       FEATURES
                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
```

ò 셤 ò g ò g ⋧ g g ð g

8

ò

q

ò

g

```
EST 02-JUL-2003
                             214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        215 CATCCCAAGACAAAAGGCGTAGCATATGCACGAACAGATATTGGCGGGCTGTACCGCCTC 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAsplleGlyGlyMetTyrArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /dev stage="mycelia"
/clone lib="T.reesei mycelial culture, Version 3 april"
/note="Wetcor: pREPBYY Site=1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                               LUD-130232 Treesei mycelial culture, Version 3 april Hypocrea jecorina cDNA clone tric029xol3, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 TyrThrTrpSerAsnValAlaIle -- - GlyGlyGlyGlyPheValAspGlyIleValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hypocrea jecorina
Eukaryota, Fungi, Ascomycota, Pezizomycotina, Sordariomycetes,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hypocreomycetidae, Hypocreales, Hypocreaceae, Hypocreal, Chases 1 to 761)

Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Broun, Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J. Kelley, A.S., Merman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M. Transcriptional regulation of biomass-degrading enzymes in filamentous fungus Trichoderma reesei
22803314
```

Thu May

g ò d ò g ઠે g 8

ò 셤 ò

```
02-JUL-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AsnWetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 580 TITICCAAGGCCTCGTCACGGCCCCTGGGACGTCCCTCCCAGACCCGAGTGATTCC 639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ThrGlyTyrGlnSerAsplleGlnGlyValValTrpValAlaPheAspLysSerSer 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
                                                                                                                                                                                                                                                              814 bp mRNA linear EST 02-JUL-:
tric074xd02 T.reesei mycelial culture, Version 3 april Hypocrea
jecorina cDNA clone tric074xd02, mRNA sequence.
CB905388
                                                                                                                                                                                                                                          6 TyrThrTrpSerAsnValAlaIle---GlyGlyGlyGlyPheValAspGlyIleValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                    ::: |||::: ||||::232 AAC---GCCGACGACGACGGATGGGATTGCTGATAATGCCGGCTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           340 GTGTATGCCGCAGTCGCATGTATGGAACAGCTGGGATCCGAGTAATGGAGCCATCATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  101 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr
                                                                                                                                                                                                                                                                                                                   25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp
                                                                                                                                                                                                                                                                                                                                                                                             45 AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal-------GlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         460 AACATGCCAGGACCCGGAGAGAGCGTCTGGCTGTCGATCCGGCCACCTCCAACATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Entaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreacea; Hypocreacea.

1 (bases 1 to 814)
Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J., Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                201 SerLeuGly-GlnAlaSerLysThrIlePheValGlyValAlaAsp 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
                                                                       747
1112
36
10
                                                                                                           Conservative:
Mismatches:
Indels:
                                                                     Length:
Matches:
                                                                                                                                                                   Gaps:
                                                                                                                                                                                                        US-09-917-376-3 (1-740) x CF876916 (1-747)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CB905388.1 GI:30120046
                                                                       7.42e-41
587.50
68.52%
51.85%
                                                                                                                                Best Local Similarity:
                                                                                                           Percent Similarity
                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                           ..
Ю
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CB905388
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      臼
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tric074xd02.b1 T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone tric074xd02, mRNA sequence.
CF876916
                                                                                                                                                                                                                                                              MetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeu 141
                                                                                                           82 TrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArg 101
                                                                                                                                  SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 121
                                                                                                                                                                                                            TyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrp 161
                                                                                                                                                                                                                                                                                                                                                                                                                                   GlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              685 GGCTACCACAGCGATCAGCAAGGACTCATGTGGGTTACGTTCGACTC-ACCAGCAGCACCACA
                                                         SerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThr 181
 275 TTCTCCGCGACTCATGGACCGCC-GTCACGGATGGGGATTGCTGATAATGCCGGCTGGCAC 333
                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreacea;
1 (bases 1 to 747)
Diener,S.E., Dankmeyer,L., Dunn-Coleman,N., Houfek,T.D.,
Mitchell,T.K., van Solingen,P., Teunissen,P.J.M., Ward,M. and
                                    62 AsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dean, R.A.
Analysis of the protein processing and secretion pathways in a factoderma reesel EST dataset
L Unpublished (2003)
Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph dean@ncsu.edu
Seq primer: IT-F1 primer.
Seq primer: IT-F1 primer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               db_xref="taxon:51453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |One="tric074xd02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CF876916.1 GI:38131598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202 LeuGlyGlnAlaSer 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     744 ACCGGNGGAGCCACG 758
                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                       142
                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 (
                                                                                                                                                                                                                                                                                                  505
                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
CF876916
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMMENT
```

160

288

9

24

44

339

80

g y

g

g

ò

```
tric082xn19.b1 T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone tric082xn19, mRNA sequence.
     181 ThrGlyTyryrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLySSerSerSer 200
                       107 AACGGTACCACCGCGACAAGGCCTCATGTGGGTTACGTTCGACTCACCCAGCATC 766
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93 TACGGCACCGGAATGACAATCTTTGGCGGCCACGATCTCACCAACTGGGAACGCGCCAC 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  377 LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note="Vector: pREPNY, Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AATGTGTCAATCCAATCACTGGCAGCAATCGAGGAATTCTCCGTCCAGGACCTGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlnIleHisIleAlaProMetValLysGlyLeuGluGluGluThrAlaValAsnAspLeuIle
                                                                                                                                                                                                                                                                                                                   Pezizomycotina; Sordariomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                Analysis of the protein processing and secretion pathways in
Trichoderma reesei EST dataset
Unpublished (2003)
Contact: Ralph A. Dean
                                                                                                                                                                                                                                                                                                                               Hypocreaceae; Hypocreales; Hypocreaceae; Hypocrea. I Chases I to 707)
Diener, S. E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D.,
Mitchell, T. K., van Solingen, P., Teunissen, P. J. M., Ward, M.
                                                                       215
                                                                                                       812
                                                                                            201 SerLeuGly-GlnAlaSerLysThrIlePheValGlyValAlaAsp
                                                                                                                                                                                                                                                                                 reesei)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           707
92
34
90
111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                               Hypocrea jecorina (anamorph: Trichoderma
Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fungal Genomics Laboratory
North Carollina State University
(Campus Box 7251, Raleigh, NC 27695,
Tel: 919-513-0020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /db_xref="taxon:51453"
/clone="tric082xn19"
                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungi; Ascomycota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-3 (1-740) x CF880713 (1-707)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Email: ralph dean@ncsu.edu
Seg primer: LT-F1 primer.
Location/Qualifiers
                                                                                                                                                                                                                                             CF880713.1 GI:38135395
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.39e-26
415.50
55.51%
40.53%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fax: 919-513-0024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
COMMENT
                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
                                                                                                                                         RESULT 5
CF880713
                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                 엄
                                                                     ò
                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                                                                                                                                                                                                            /dev_stage="mycelia" /clone lib="T.reesei mycelial culture, Version 3 april" /clone lib="T.reesei mycelial : Not I/Sal I; Mycelial /note="Vector: pREP3Y; Site_I: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TyrThrTrpSerAsnValAlalle---GlyGlyGlyGlyPheValAspGlyIleValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           356 CACAACTGG-------GGCATCGACGCTGTTGCGCTTCATCCGCAGGACGATCAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal------GlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M. Transcriptional regulation of biomass-degrading enzymes in filamentous fungus Trichoderma resesi J. Biol. Chem. 278 (34), 31988-31997 (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       814
112
36
58
58
4
                                                                                                                                         94304, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                            /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                         S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                    Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, Cl
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
                                                                                                                                                                                                                                                                                                            db xref="taxon:51453"
clone="tric074xd02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-3 (1-740) x CB905388 (1-814)
                                                                                                                                                                                                          Seq primer: LT-FI primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8.53e-41
587.50
68.52%
51.85%
                                                                                                                                                                                                                                               1. .814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                     12788920
                                                                     22803314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          547
                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                   MEDLINE
PUBMED
                                                   JOURNAL
                                                                                                                                                                                                                             FEATURES
                                                                                                     COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
```

g

ò

g

ò

셤

ò

g

ò

g

ò

g

ð

g

ò

셤 ò

ò

. .

92

416

287 456 347 473 407 493 467 513 512 533

```
tric029xol3.bll T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone tric029xol3, mRNA sequence.
CF882065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  534 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla 553
/note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           168 ACCGCCGCACGAACAG------GTGGCCATCTCGTCCGACGGCGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TACGGCACCGGAATGACAATCTTTGGCGGCCACGATCTCACCAACTGGGACACGCGCCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Accregacidacidaciacidacidacidacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas accregacidas acceptadas accregacidas accregacidas acceptadas accregacidas acceptadas accregacidas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptadas acceptad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          514 AsnrrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    513 ACGIGGAGCATCGACTACGCGGCCGACACGTCCATGAACGGCGCGCACGGTGGCCTATICG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ecceaceaceacarcercreerceacecercescercesec-----erecaeee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             554 ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        624 regeagriceassecascerriscerereseasecriscesesecécsesecécsereares
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ala------AspValThrAlaValProSerThrIlePheThrSerProValPheThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
                                                                                                                                                                                            782
92
34
90
111
                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                      (1-782)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        744 AAGGACACCGGCAGCATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    594 ThrAspGlyGlyValThrPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CF882065.1 GI:38136747
                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-3 (1-740) x CB907625
                                                                                                                                                                                  8.7e-26
415.50
55.51%
40.53%
10.29%
                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     494
                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
DEFINITION
                                                                                                                                                                                                      ..
9
                                                                                                                                                                  Alignment Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CF882065
                                                                                                    DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               유
장
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CB907625 Treesei mycelial culture, Version 3 april Hypocrea jecorina cDNA clone tric082xn19, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             548
                                                                                                                                                                  332
                                                                                                                                                                                                                                                                       513
                                                                                                                                                                                                                                                                                                                                                                                                                         437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          554 ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                668
                                                                                                                                                                                                                                474 ThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone="tric082xn19"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
                                                                                                                                                                                                                                                                                                                                                                                            393 ACCGCCGCACGAACAG------GTGGCCATCTCGTCCGACGGCGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hypocreonecee; Hyp
   TCGCAGTTCCAGGGCAGCTTGCCTCCGTCTCGAGCCTGCCCGCGGGGCGCCGTCATCGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSer
                                                                                                    Ala-----AspvalThrAlaValProSerThrIlePheThrSerProValPheThr
                                                                                                                                                                                                                                                                                                                                                        494 PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   514 AsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 438 ACGTGGAGCATCGACTACGCGGCCGACACGTCCATGAACGGCGGCGCGCGGGGCCTATTCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               534 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         rcegacaagaagaccaacagcgrcrrcracecceccrccegarcgaccrrrracgrcage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CA 94304, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
22803314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              organism="Hypocrea jecorina"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric082xn19"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              925 Page Mill Road, Palo Alto,
Tel: (650) 846-7635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fax: (650) 621-7817
Email: Pforeman@genencor.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seq primer: LT-FI primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAGGACACCGGCAGCAGCTTC 689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThraspGlyGlyValThrPhe 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Contact: Pamela K. Foreman
Genencor Intl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CB907625.1 GI:30122283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12788920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CB907625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EST.
                                                                                                                                                                                                                                                                                             333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               699
                                          213
                                                                                                    457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
CB907625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                    ò
```

```
255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                    RESULT 8
CF868882
LOCUS
                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98 TITICAIGGAAGAACGICAAGCTCGGCGGCGGCGGCGTTCGTCCCCGGCATCATCTTC 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 101 ArgserSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'dev_stage="mycelia"
|clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ::: |||::: |||::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: |||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || ::: || 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GIGTATGCCGCACGCCGCATGTATACGAACACCTGAGATCCTACTAATGGACCCATCATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 TyrThrTrpSerAsnValAlaIle -- - GlyGlyGlyGlyPheValAspGlyIleValPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            45 AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal------GlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        386 cccricárcada de concescio de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de con
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea. 1 (bases 1 to 693) Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D., Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M. and
                                                                                                                                                                                                                                                    Analysis of the protein processing and secretion pathways in
Trichoderma reese! EST dataset
Unpublished (2003)
Contact: Ralph A. Dean
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuTyrPheGlyAlaProSerGlyLysGlyLeu-TrpArgSerThr 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arcracrececrecaeaaacaecrecrecaeaacrecrecaeaacrerace 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    693
79
24
43
10
4
                                                                                                                                                                                                                                                                                                                                                                                                                      Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph deamencsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="Hypocrea jecorina"
/mol type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         db_xref="taxon:51453"
clone="tric029xo13"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-3 (1-740) x CF882065 (1-693)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                385.50
66.03%
50.64%
9.55%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                . 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                      Dean, R.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                           REFERENCE
                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                   COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
tricol6xm14.bl T.reesei mycelial culture, Version 6 October 2003 Hypocrea jecorina cDNA clone tricol6xm14, mRNA sequence. CF868882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 464 SerThrilePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /clone="tric016xm14"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            195 CAGACGGTCTGGGCAACGCCCACATGGGCCACCTCGACGAGGGTCGACCTACGCCGGGAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            407 AsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         427 LeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       447 AlaLeuGlyAspLeuGlyGlyPheThrHisAla-----AspValThrAlaValPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              484 AsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TCGGTCAAGAGCGTCCGCGTCGGCAACACCCGCCGGCACGG-------
                                                                                                                                                                                                Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

    (bases 1 to 751)

                                                                                                                                                                                                                                                                                                                                   Analysis of the protein processing and secretion pathways in
Trichoderma reesei EST dataset
Unpublished (2003)
Contact: Ralph A. Dean
Fungal Genomics Laboratory
                                                                                                                                                                                                                                                                 Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D.,
Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M.
                                                                                                                                                      reesei)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  751
88
40
103
17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                 Hypocrea jecorina (anamorph: Trichoderma
Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                     North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-3 (1-740) x CF868882 (1-751)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fax: 919-513-0024
Email: ralph dean@ncsu.edu
Seq primer: IT-FI primer.
Location/Qualifiers
                                                                                                      CF868882.1 GI:38123564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.29e-21
361.50
51.61%
35.48%
8.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
                                                                                                                                                                                                                                                                                                                     Dean, R.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
```

9 à 셤 ઠે 셤 à 셤 8 d 충 g ò g ð 셤

```
linear EST 31-OCT-2003
Version 6 October 2003
mRNA sequence.
                                                                                                                                                                                                                          426
                                                                                                                                                                                                                                                          126
                                                                                                                                                                                                                                                                                                  446
                                                                                                                                                                                                                                                                                                                    127 ATCGAGGAATTCTCCGTCCAGGACCTGCCCTCTGCACCGGGGGAAGCGAGCTATTGGCC 186
                                                                                                                                                                                                                                                                                                                                                                         463
                                                                                                                                                                                                                                                                                                                                                                                                                                                  483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              584 AlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProVal 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       691 ACCACCGCGGGCACGTTGTATGTCTCCACCGACGTCGCCATATTCCGCTCCACACACTCG 750
                                                                                                                                                                                                                                                                                                                                                                                            524 ValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               522 GlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGly
                                                                                                                                                                                                                                                  LeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSer
                                                                                                                                                                                                                                                                                                                                                                       AlaLeuGlyAspLeuGlyGlyPheThrHisAla-----AspValThrAlaValPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AsnProSerIleIleValArgAlaGlySerPheAspProSerSerGinProAsnAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               504 HisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     172 GCCTCGTCCGGC-----GTGCAGCGCTCGCAGTTCCAGGGCAGCTTTGCCTCGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        564 GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              523 TCGAGCCTGCCCGCGGGGCGCCGTCATCGCCTCGGACAAGAAGACCAACAACAGGGGTCTTCTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         504 AlaAlaGlyLeuProSerSerGlyAlaValGlyValMet-----PheHisAlaValPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407 AsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       803
88
40
1103
117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 trico06xe08.bl T.reesei mycelial culture,
Hypocrea jecorina cDNA clone trico06xe08,
CF866388
CF866388.1 GI:38121014
                                                                        Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EST.
Hypocrea jecorina (anamorph: Trichoderma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mRNA
                                                       Length:
                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    738 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          542 GlySerSerTrpSerAlaIleThr 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 x CB898982 (1-803)
                                                       78e-21
                                                                      361.50
51.61%
35.48%
8.96%
                                                                                                                                                                                     US-09-917-376-3 (1-740)
                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CF866388
                                        Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   544
                                                                                                                                                                                                                                                                                                                                                                                                                                                    464
                                                                                                                                                                                                                                                                                                  427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               751
                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CF866388
LOCUS
DEFINITION
                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
KEYWORDS
SOURCE
                                      Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                  ठ
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ଚ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EST 02-JUL-2003
                                                                                                                                                                                                                                                                                                                                                                                                               638
                                                                                                                                                                                                                                                              530
                                                                                                                                                   563
                                                                                                                                                                                     470
                                                                                                                                                                                                                          583
                                                                                                                                                                                                                                                                                                  603
                                                                                                                                                                                                                                                                                                                                       581
                                                                                                                                                                                                                                                                                                                                                                           621
                                                                                                                                                                                                                                                                                                                                                                                                                                                    641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone="tricol6xm14"
/dev stage="mycelia"
/clone lib="T.reesei mycelial culture, Version 3 april"
/note="Verctor: pREP3Y; Site_1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CB898982 803 bp mRNA linear EST 02-JUL-1
tric016xm14 T.reesei mycelial culture, Version 3 april Hypocrea
jecorina cDNA clone tric016xm14, mRNA sequence.
HisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly
                                                                                                                                                                                                                                                                                                  584 AlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProVal
                      ---GTGGCCATCTCGTCCGACGCGGCGACGTGGAGCATCGACTACGCGGCCGACACG
                                                                      524 ValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro
                                                                                               TCCATGAACGGCGCACGGTGGCCTATTCGGCCGACGCGACACGATCCTCTGGTCGACC
                                                                                                                                             544 GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSer
                                                                                                                                                                                     GCCTCGTCCGGC-----GTGCAGCGCTCGCAGTTCCAGGCCAGCTTTGCCTCCGTC
                                                                                                                                                                                                                          564 GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyr
                                                                                                                                                                                                                                                              471 TCGAGCCTGCCCGCGGCGCCGTCATCGCCTCGGACAAGAGACCAACAGCGTCTTCTAC
                                                                                                                                                                                                                                                                                                                                     531 GCCGGCTCCGGATCGACCTTTTACGTCAGCAAGGACACCGGCAGCATC-----
                                                                                                                                                                                                                                                                                                                                                                           604 AlaAlaGlyLeuProSerSerGlyAlaValGlyValMet-----PheHisAlaValPro
                                                                                                                                                                                                                                                                                                                                                                                                             582 ACGCGCGCG---CCCAAGCTGGGCAGCGCAGGGACGATCCGGGATATCGCTGCTCACCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pezizomycotina; Sordariomycetes,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycete Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

1 (bases 1 to Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J. Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M. Transcriptional regulation of biomass-degrading enzymes in t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CA 94304, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
22803314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genencor Intl.
925 Page Mill Road, Palo Alto,
Tel: (650) 846-7635
Fax: (650) 621-7817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlySerSerTrpSerAlalleThr 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGCACGACCTTTGGCCAAGTCTCC 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Email: Pforeman@genencor.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seq primer: LT-Fl primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Contact: Pamela K. Foreman
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CB898982.1 GI:30113640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CB898982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12788920
                                                                                                             360
                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CB898982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
```

```
EST 02-JUL-2003
                                                                                                                                       617 ---PheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 635
 398 GGCAGCTTTGCCTCCGTCTCGAGCCTGCCCGCGGGCGCCGTCATCGCCTCGGACAAGAAG 457
                                                                                                                                                                                                                                                                      636 TyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAla 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /db_xref="taxon:51453"
/clone="tric006xe08"
/dev stage="unvealia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/note="Vector: pREP3Y; Site_1: Not I/Sal 1; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CB896354 1.reesei mycelial culture, Version 3 april Hypocrea jecorina cDNA clone tric006xe08, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

I (bases 1 to 794)
Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J., Relley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J. and Mard, M.

Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
                                                           598 ValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMet - - -
                                     578 AsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGly
                                                                                                                                                                                                                                                                                                                                             655 ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr 668
                                                                                                                                                                                                                                                                                                                                                                     Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               794
89
43
102
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. .794
/organia="Hypocrea jecorina"
/mol type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tel: (550) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seq primer: LT-FI primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Pamela K. Foreman
Genencor Intl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CB896354.1 GI:30111012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.32e-19
338.00
51.97%
35.04%
8.37%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12788920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                             CB896354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                 셤
                                                                                                         ð
                                                                                                                                               임
                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                               ò
                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                422 ProMetValLys---GlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSer 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                242 CAACAG-------GTGGCCATCTCGTCCGACGGCGGCGCGCGTGGAGCATC 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :::|||:::
147 ACGATCCTCTGGTCGGCCTCGTCCGGC-----GTGCAGCGCTCGCAGTTCCAG 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AsnSerTrpalaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 287 GACTACGCGGCCGACACGTCCATGAACGGCGGCGACGGTGGCCTATTCGGCCGACGGCGAC 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAla-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      518 GlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               538 ArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGly
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
1 (bases 1 to 738)
                                                                     Diener, S.E., Dankmeyer, L., Dunn-Coleman, N., Houfek, T.D., Mitchell, T.K., van Solingen, P., Teunissen, P.J.M., Ward, M. and
                                                                                                                         Death, A.A.

Trichoderma reesei EST dataset
Unpublished (2003)

Contact: Ralph A. Dean

Fungal Genomics Laboratory

North Carolina State University

Campus Box 7251, Raleigh, NC 27695, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         738
89
43
102
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                 /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                                                                                                                                                                                         db_xref="taxon:51453"
clone="tric006xe08"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-3 (1-740) x CF866388 (1-738)
                                                                                                                                                                                                                                                                                  Fax: 919-513-0024
Email: ralph dean@ncsu.edu
Seg primer: LT-F1 primer.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.73e-19
338.00
51.97%
35.04%
8.37%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                 Dean, R.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                      JOURNAL
                                                     REFERENCE
                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                 TITLE
                                                                                                                                                                                        COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DRIGIN
```

ò d g

ò

a

ò

d

ò

셤

ò

g ò 쉽 ò

à

셤 8 g δ g 8 유 ð g ò g ò a ò 원 ò 셤 à

ઢ

```
/.dev stage="Mycelium"
/lab_host="E. coli"
/lab_host="E. coli"
/lone lib="Subtracted Mycelial"
/note="Vector: pBlueScript SK (-); Site_1: EcoRI; Site_2:
XhoI; 2* sucrose for 24 hours. CDNA directionally cloned
into pBluescript SK(-) using the Unit_ARP KR vector system
(Stratagene, La Jolla, CA). Previously identified highly
expressed clones were subtracted from this library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       56 CTCGGCTGGATGGATCGAATCCCTCGAGATCAACCCCCTCGACGACGACCATTGGCTCTAC 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       398 GlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGln 417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116 GCCACCGGTCTCACTGTTTACGGCGGCCACCTGACCAGAGTGGGACACCGTCCACAAC 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsn 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    116 TCCGGCGCGCAGCAG------GTGCCCGTATCCTCGGACGCGCGCTCG 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175 GlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPhe 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLys 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyr 397
   ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            138 ProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThr----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      356 rececegacerceachacecegeaacaaccececeaacereerecaceareececec
   Expressed sequences from conidial, mycelial, and sexual stages
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    418 IleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              176 GTGACCATCCAGTCCTTGGCCGTTGGCATCGAAGAAATGGCTGTCCTCGGTCTCGCCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      456 ---HisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           546
72
30
75
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /moi_type="mRNA"
fstrain="74-OR23-IV A (FGSC 2489)"
/db ref="taxon:5141"
/clone="SM3H7"
                                                                                                                                                                                                                                                 USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                 87131,

    .546
    /organism="Neurospora crassa"

                                oora crassa
Genet. Biol. 21, 348-363 (1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                             Contact: Natvig,D.O./Nelson,M.A.
Department of Biology
University of New Mexico
Tatetter Hall, Albuquerque, NM 87
Tel: 505 277 3411
Fax: 505 277 0304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /sex="Mating type A"
/tissue_type="Mycelium"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (1-546)
                                                                                                                                                                                                                                                                                                                                        ngp@biology.unm.edu.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-3 (1-740) x BF072664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3.29e-17
314.50
54.26%
38.30%
                                   Neurospora
Fungal Gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                          97435549
                                                                                                                      9290248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           No.:
                                                         JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment
                                                                                                                                                                                                                                                                                                                                                                        FEATURES
TITLE
                                                                                                                                                       COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF072664 Subtracted Mycelial Neurospora crassa cDNA clone SM3H7 5'smilar to avicelase III, Aspergillus aculeatus, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Fund; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

1 (Bases I to 546)

Nelson, M.A., Kang, S., Braun, E.L., Crawford, M.E., Dolan, P.L., Leonard, P.M., Mitchell, J., Armijo, A.M., Bean, L., Blueyes, E., Cushing, T., Errett, A., Fleharty, M., Gorman, M., Judson, K., Miller, R., Ortega, J., Pavlova, I., Perea, J., Todisco, S., and Natvig, D.O.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  616
                                                                                          457
                                                                                                                                                                                                                    477
                                                                                                                                                                                                                                                                                                                                                 518 GlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySer 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         617 ---PheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProMetValLys --- GlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer 440
                                                                                                                                                                                                                                                                                                                                                                                                                                        478 ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSer 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gricialcia de de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  498 SerGinProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGln 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   343 GACTACGCGGCCGACACGTCCATGAACGGCGCACGGTGGCCTATTCGGCCGACGGCGAC 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   538 ArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGly 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              558 AsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            574 AGCAGCTTC-----ACGCGCGGG---CCCAAGCTGGGCAGCGCAGGGACGATCCGG 621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   636 TyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAla 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GGCAGCTTTGCCTCCGTCTCGAGCCTGCCGCGGGGCGCCGTCATCGCCTCGGACAAGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               598 ValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMet---
                                                                                                                                                                                    GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAla-----
                                                                                                                                                                                                                                                                                                           AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        578 AsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  514 ACCAACAGCGTCTTCTACGCCGGCTCCGGATCGACCTTTTACGTCAGCAAGGACACCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TACCAGATCGCCCTGGGTGTGGGCTCA---GGCTCGAACTGG 780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       655 ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr 668
JS-09-917-376-3 (1-740) x CB896354 (1-794)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BF072664.1 GI:10866169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Neurospora crassa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    742
                                                                                                                          28
                                                                                                                                                                                    441
                                                                                                                                                                                                                                                 118
                                                                                                                                                                                                                                                                                                              458
                                                                                                                                                                                                                                                                                                                                                                           178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF072664
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
```

8

ਨੇ 셤 ò a à

RESULT 13

셤 ò

셤

BU639045

ACCESSION

```
linear EST 31-OCT-2003
Version 6 October 2003
mRNA sequence.
                                                                                                                                                                                                                176 CATTGGCTTTACGGAACTGGCCTCTCGATCTTTGGGGGACATGACTTGACCAAATGGGAC 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296 GAGATGGCTTCAGCGCCCGGAGGCTCCGAACTACTGGCGCGCAATTGGAGACATTACTGGA 355
                                                                                                                                                                                         354 ValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProVal 373
                                                                                                                                                                                                                                                                          374 ProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAsp 393
                                                                                                                                                                                                                                                                                                                                                            394 ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAsp 413
                                                                                                                                                                                                                                                                                                                                                                                                                                            414 SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsn 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   236 ACGCTTCATAACATATCGATTCACTCTTTGGCCGATGGTGTCGAAGAAACTGCAGTTCTG 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            434 AspleulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGlyGly 453
                                                                                                                                 5 GGGCTACATGGTCGAGACTTTGGGAGTGGACGAGCTATCCCGACATGAACCAGTACTAC 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="Vector: pREPNY, Site 1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Analysis of the protein processing and secretion pathways in a Trichoderma resest EST dataset Unpublished (2003)
Contact: Ralph A. Dean Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreacea; Hypocrea. 1 (bases 1 to 713) Diener, S. B., Dankmeyer, L., Dunn-Coleman, N., Houfek, T. D., Mitchell, T. K., van Solingen, P., Teunissen, P.J. M., Ward, M. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hypocrea jecorina (anamorph: Trichoderma reesei)
Hypocrea jecorina
  mH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tric039xk06.bl T.reesei mycelial culture,
Hypocrea jecorina cDNA clone tric039xk06,
CF875815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      713 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     db xref="taxon:51453"
                                                                US-09-917-376-3 (1-740) x BU639045 (1-389)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         clone="tric039xk06"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Email: ralph dean@ncsu.edu
Seg primer: LT-Fl primer
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CF875815.1 GI:38130497
      7.59%
13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fax: 919-513-0024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           454 PheThrHis 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |||||||:::
356 TTCACATAC 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dean, R.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CF875815
      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
CF875815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                         g
                                                                                                                                                                                       ò
                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                        ⋧
                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BU639045 389 bp mRNA linear EST 06-MAY-2003 mgcw011xD18f.b RCW Lambda Zap Express Library Magnaporthe grisea cDNA clone mgcw011xD18 5', mRNA sequence.
      534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Email: d-ebbole@tamu.edu
Chromatcogram file of this sequence is available, see contact
person;Best nr hit (April. 22, 2003) gb|EAA29333.1| hypothetical
protein [Neurospora crassa] 206 4e-53
                                                                                                                                                                                                                                                                                                                                                                                                                                   Ebbole, D.J., Yuan, J., Thomas, T.L., Bobrowicz, P., Lu, G., Bhatterai, K. and Dean, R.A.
Expressed sequence tags from the rice blast fungus, Magnaporthe
TrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaSerAla
                                           461 TGGCATGCTCACAACGGCACCGACACCAAGAGCAGCGGCACGGTCGCCTACTCTGCC
                                                                                                                                                                                                                                                                                                                                                                              Pezizomycotina; Sordariomycetes;
                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    grisea
Unpublished (2002)
On Sep 30, 2002 this sequence version replaced gi:23351371.
Contact: Ebbole DJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Department of Plant Pathology & Microbiology
Texas A&M University
Peterson. Bldg MS2132, College Station, TX 77843-2132,
Tel: 979 845 4831
Pax: 979 845 6483
                                                                                                                                                                                                                                                                                                                                  Magnaporthe grisea (anamorph: Pyricularia grisea)
Magnaporthe grisea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /organism="Magnaporthe grisea"
/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sex="Mat1-1 hermaphrodite"
tissue_type="Mycelium"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BACKWARD: T7 primer
Plate: mgcw011 row: D column: 18
Seg primer: T1.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db_xref="taxon:148305"
/clone="mgcw011xD18"
                                                                                 535 AspGlySerArgPheValTrpAla 542
                                                                                                         521 GACGCCGACACCATCGTCTGGTCT 544
                                                                                                                                                                                                                                                                                            BU639045.2 GI:30397224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    strain="CP987"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9.49e-17
306.50
66.67%
44.72%
                                                                                                                                                                                                                                                                                                                                                                                                                    (bases 1 to 389)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FORWARD: T3 primer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
  515
                                                                                                                                                                                                        LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
```

FEATURES

ORIGIN

JOURNAL COMMENT

TITLE

REFERENCE AUTHORS

8 요 8 g ò 원 ò g ò g ò g à 원 ò 셤 ò 8 \$ 셤 ð 셤 8 셤 ò g

```
537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AcGATCCTCTGGTCGACCGCCTCGTCCGGC-----GTGCAGCGCTCGCAGTTCCAG 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            474
                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone="tric033xk06"
/dev_stage="mycelia"
/clone lib="T.reesei mycelial culture, Version 3 april"
/note="Vector: pREP3Y, Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GGCAGCTTTGCCTCCGACCTGCCCCGCGGGCGCCGTCATCGCCTCGGACAAGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       458 AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GTCGACTACGCCGGGAACTCGGTCAAGAGCGTCGTCGCGCTCGCGTCGCGCAACACGCCGGCACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GACTACGCGCCCGACACGTCCATGAACGCCGCACGCTGGCCTATTCGGCCGACGCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Accaacagegrerreracgecegerecageregaeerreracgaeercageaegaeegee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMet ---
                                                                                                                                         the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGly
                   Sordariomycetes
                                                            Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J. Kellery, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M. Transcriptional regulation of biomass-degrading enzymes in tilamentous fungus Trichoderma ressei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
                                  Hypocrea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            and Nitrogen sources and concentrations.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                782
81
40
106
21
Hypocrea jecorina
Bukaryota; Fungi; Ascomycota; Pezizomycotina;
Hypocreomycetidae; Hypocreales; Hypocreaceae;
I to 782)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                94304,
                                                                                                                                                                                                                                                                                                                                                                               /organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                            Genencor Intl.
925 Page Mill Road, Palo Alto,
721: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
                                                                                                                                                                                                                                                                                                                                                                                                                                      'db_xref="taxon:51453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (1-782)
                                                                                                                                                                                                                                                                                                                                Seq primer: LT-FI primer.
Location/Qualifiers
                                                                                                                                                                                                                            Contact: Pamela K. Foreman
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     x CB904767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.65e-14
286.50
48.79%
32.66%
7.10%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-3 (1-740)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAACAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
 ORGANISM
                                                  REFERENCE
AUTHORS
                                                                                                                                                                                        MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment
                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EST 02-JUL-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      616
                                                                                                                                                                                                                                                                                                GTCGACTACGCCGGGAACTCGGTCAAGAGCGTCGTCGCGTCGGCAACACCGCCGGCACG 129
                                                                                                                                                                                                                                                                                                                                                    517
                                                                                                                                                                                                                                                                                                                                                                                                                      537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCAACAGCGTCTTCTACGCCGGCTCCGGATCGACCTTTTACGTCAGCAAGGACACCGGC 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---PheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   692
                                                                                                                                                                                                                                                                              497
                                                                                                                                                                                                                                                                                                                                                                      234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGly 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     681
                                                                                                                                                                                                                                            69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CB904767
TR2 bp mRNA linear EST 02-JUL-2 tric039xk06 T.reesei mycelial culture, Version 3 april Hypocrea jecorina cDNA clone tric039xk06, mRNA sequence.
CB904767 GI:30119425
EST.
Hypocrea jecorina (anamorph: Trichoderma reesei)
                                                                                                                                                                                                                                                                                                                                                SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                      GACTACGCGGCCGACACGTCCATGAACGGCGGCGCGTGGCCTATTCGGCCGACGGCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACGATCCTCTGGTCGACCGCCTCCGGC-----GTGCAGCGCTCGCAGTTCCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMet ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGCACC-----GGCCCGTCAGGGCTCGCCTCTACGCCAGTGGAGACAGCGGCGTCCC
                                                                                                                                                                                                          AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer
                                                                                                                                                                                                                              TICCGCTCCACAGACTCGGGCACGACCTTGGCCAAGTCTCCACGGCCTGACGACACACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GlyThr11eGlyGlyValThrGlyAla-----TyrArgSerAspAspCysGlyThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                    GlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               574 TACCAGATCGCCCTGGGTGTGGGCTCA---GGCTCGAACTGG---AACCTGTATGCCTTC
                                                                                                                                                                                                                                                                              ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValVal
                                                  713
81
440
106
21
                                              Length:
Matches:
Conservative:
Mismatches:
                                                                                                                      Indels:
                                                                                                                                        Gaps:
                                                                                                                                                                        (1-713)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TrpValLeuIleAsnAspAspGln 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     redacedearireaecereceáe
                                                                                                                                                                        US-09-917-376-3 (1-740) x CF875815
                                                  1.42e-14
286.50
48.79%
32.66%
7.10%
                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                            10
                                                                                                                                                                                                                                                                              478
                                                                                                                                                                                                                                                                                                                20
                                                                                                                                                                                                                                                                                                                                                  861
                                                                                                                                                                                                                                                                                                                                                                                                                    518
                                                                                                                                                                                                                                                                                                                                                                                                                                                      175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CB904767
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
ORIGIN
```

```
475 AGCAGCTTC-----ACGCGCGGG---CCCAAGCTGGGCAGGGAGGAACGATCCGG 522
                                  617 ---PheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 635
                                                655 ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValVal 674
                                                                                                                                 643 TACCAGATCGCCCTGGGTGTGGGCTCA---GGCTCGAACTGG---AACCTGTATGCCTTC 696
                                                                                                                                                                                                     693 TrpValLeulleAsnAspAspGln 700
                                                                                                                                                                                                                         751 TGGACGGACATTCAGGGCTCCCAG 774
                                ઠે
                                                   g
                                                                        ઠે
                                                                                             g
                                                                                                              ò
                                                                                                                                   QQ
                                                                                                                                                          ઠે
                                                                                                                                                                              gg
                                                                                                                                                                                                   à
```

Search completed: May 12, 2004, 11:38:55 Job time : 6964.51 secs

5730, Ap 3, Appli 2, Appli

Sequence

Sequence

29, Appl 7, Appli 1, Appli 57, Appl

Sequence Sequence

Sequence Sequence Sequence

Sequence Sequence

Sequence Sequence

Title: Perfect score:

Run on:

Sequence:

Scoring table:

Searched:

```
### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:

### APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2993 AACGGCGTCGTCATCAGCGGCACCGCCGAGGCCGTGCCACCGTGACCCTCACCGATGCC 2934
                                                                                                                                   Sequence 2642. Ap
Sequence 17. Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 14. Appli
Sequence 1420, Ap
Patent No. 5457037
Sequence 75, Appli
Sequence 2, Appli
Sequence 15371, A
Sequence 15403, A
Sequence 15403, A
Sequence 15403, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20 AspGly1leValPheAsnGluGlyAla---ProGlyIleLeuTyrValArgThrAspIle 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence Sequence Sequence 3
                                                                                                            Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
US-07-998-931-4
US-09-679-279-1
US-09-479-179-1
US-09-408-020-3
US-09-408-020-3
US-09-408-020-2
US-09-408-020-2
US-09-252-991A-9781
US-08-228-662-1
US-08-852-219-1
US-08-852-219-1
US-09-023-655-1420
S-557037-1
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-991A-15371
US-09-252-291A-15371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-3 (1-740) x US-09-252-991A-13873 (1-3129)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-252-991A-5391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.76e-11
227.50
31.30%
22.15%
5.64%
  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-252-991A-13873
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                             163
160.5
159.5
159.5
159.5
157.5
156.5
156.1
155.1
155.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/CQII_1/GSPTO spool/1376/runat 11052004 114536 28265/app_query.fasta_1.1429
-Q=/CQII_1/GSPTO spool/1376/runat 11052004 114536 28265/app_query.fasta_1.1429
-DB=ISSUG_APTO spool/1376/runat 11052004 114536 28265/app_query.fasta_1.1429
-DB=ISSUG_APTO spool/1376/serata -END=-1.mARTRIX=blosum62 -TRANS=bluman40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-USTF=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-USTF=450509917376 @CGN 1 1 116 @runat 11052004 114536 28265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DBV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGRPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 13873, A
Sequence 13656, A
Sequence 13774, A
Sequence 13773, A
Sequence 17, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1377. Ap
Sequence 39, Appl
Sequence 2, Appli
Sequence 1, Appli
Sequence 13875, A
Sequence 4, Appli
                                                                                                                                                                                                                   (without alignments)
2222.227 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                      1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                          May 12, 2004, 04:43:36 ; Search time 184.798 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1: /cgn2 6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
                              GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-252-991A-13873
US-09-252-991A-13774
US-09-252-991A-13774
US-09-252-991A-13773
3 US-09-103-840A-1
3 US-09-103-840A-2
US-09-433-702B-39
3 US-09-103-840A-1
3 US-09-103-840A-1
3 US-09-103-840A-1
4 US-09-252-991A-13875
US-07-706-699-4
                                                                                                                                        OM protein - nucleic search, using frame_plus_p2n mode]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          682709 segs, 277475446 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                       0.5
0.5
0.0
                                                                                                                                                                                                                                                                                                                                                                                                                               Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Issued_Patents_NA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5.6 3129 4 U

5.6 8211 4 U

5.0 4138 4 U

6.8 4411529 3

4.6 4403765 3

4.2 21165 4 U

4.2 21165 4 U

4.2 4403765 3

4.2 441329 3

4.1 4411529 3

4.1 4411529 4

4.0 5741 1 U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length DB
                                                                                                                                                                                                                                                                                                    US-09-917-376-3
                                                       Copyright
                                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
```

Jatabase :

227.5 227.5 203.2 201.5 192.1 181.5 170.5 164.1 163.1

Result

Score		#S VALGIYMETIYINIASINETIIIDASPIRETIYASINETIIIDASPIRETIYINIATIELEGUARIA 101 101 102 102 103 103 104 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105 105	SerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGly :::	3967 GTCAACGCCGTGGCCCAGGACCTGCGGGCAATACCGGCCCGCAGGGCAGCACTACCGTG 4026 131 LeualavalaspProAsn	4087 ACCGCGAGCCGGGCAGCACCCTTGACCGACGGCAACCGCATCGGCAG 4146 153 ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhe 167 1417 ACCACGCGCGAGCGGCAACTGGAGCTTCAGGCCGCCGGCCG	ProAs;	4252GCTCCCCGCTACCACGACGGGAITCCICGCIGCCGICGAICCCGCAGGIGGAI 4909 205AlaserLysThrilePheValGlyValAlaAspProAsnAsnProValPhe 221 ::::: :::::::::::::::::::::::::::::	222	233 ValProglyAlaPro 237
Pred. Score Score Best 1 Duery US-09	6 6 6 6	6 6 6 6	& 8 &	9 & 9 & 8	9 & 8	& 8 & :	a & a	S S	Oy Dp
Db 918 GAGCTTCACGCCCGCCACGCCCAACGGCTCCGTGATCAATGCCCTGGCCCAGGA 859 Qy 575	Db 684 CAGCGGCACTTCACTCCGGGCACA-CCGCTGACCAACGGCGTGATCAATG 626 Qy 630 laAlaSerSerGlyLeuTyrHisSerThrAnGlyGlySerSerTrpSerAla1 648	Qy 668 yrProAlaValPheValValGlyThrIleGlyGly	681 hrdlyalafyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnH 681 hrdlyalafyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnH 681 hrdlyalafyrArgSerAspAspCccGccGGCAATACCAGCGGTCCGGCCAGC 681 hrdlyalafyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArg-Val		Qy 739 SerGly 740 Db 215 ATCGGC 210 RESULT 2 115-09-252-991A-13656	OD CO	WT APPLICATION WT FILING DATE: APPLICATION NU FILING DATE: APPLICATION NU	; PRIOR FILING DATE: 1998-07-27 ; NUMBER OF SEQ ID NOS: 33142 ; LENGTH: 8211	; TYPE: DNA ; ORGANISM: Pseudomonas aeruginosa US-09-252-991A-13656 Alignment Scores:

5472 CAATACCAGCGGCCCGACCACCACGGTGGACGCGGGCCCCGGCCACCCCGGTGGT 530	DD 5592 CACCGACGGCGACCCGATCGGCCAGGCCGACGCCGACGCCACCTG 5648 QY 560 palaalaSerGlnGlyValProAlaAsnalaGlnIleArgSer	Db 5769 CCCGGTGCTCGATCCGAĞCĞTÄCĞGTGATCAGCGGTACCGCCGAAGCCGGGCCAC 5828 Qy 594ThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeu-ProSerS 610		680 6122 681 6182 701 723	Oy 721 TyrileGlyThrAsnGlyArgGlyIleValTyrGlyAspileGlyGlyAlapro 738	Sequence 13774, Application US/09252991A Fatent No. 6551795 GENERAL INFORMATION: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS FILE REFERENCE: 107196.136 CURRENT APPLICATION NUMBER: US/09/252,991A
238	4606 CTCACCÓATGGCGGCAACCCGATCGGCCAGCCGCCGACGGCGACGGCACTGG 4665 283 ThrArglleSerProValPro	4786 CCGGTGATCAATCCGAGCAACGGAGTCGTGTACGGGGGAGCCGGAGGCCACG 4845 325 ThrileIlePheArgSerThrAspGlyGly		408 pLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGl 428	461 aValProSerThrIlePheThrSerProValPheThrThrGlyThrSerVal 478	5373 CCAGGCC

8 & 8 &

8 6 6

g

8 S

```
B
                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                      셤
                                                                                           셤
                                                                                                                        ઠે
                                                                                                                                                                             ò
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                             2438 ACGIGGIGGCCAGGGAIGCCGCCGGCAACAGCAGTCCGCCGGCCAGCGTIA----- 2388
                                                                                                                                                                                                                                                                                                                             GCCAACGGCAACTGGTCTTTCACGCCGTCCACCCCGCTGCCGGACG--GTACCGTGGTCA 2439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----- 2290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------GACGCAGCGCAACTGG------ 2236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2235 ---ACCTTCACCCCGAGCAGCGTTGCCCAAC---GGCACGGTGGTGAACGCCACGGCT 2182
                                                                                                                                                                                                                                                                                                                                                                                                                                      ---cceredar-eccercecceccaccaccaccarcarcarcaaccaacaacaaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GCCACGCCAGTGGTCAACCCGAGCAACGCCACCTCAGCGGCACCGCCGAGCCGGGC 2062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GCCACCGTGACCCTGGCCGATGGCCAATGCCATCGGGCAGGTCACCGCCGATGGC 2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGCGGCAACTGGAGCTTCACTCCGACCACGCCCTAACGGCACCGTGGTCAACGCC 1942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1881 GCCCCGGCCACGCCAGTGATCAACCCCAACGGCACCACGCTCAGCGCCACCCCGAG 1822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----AsnAsnProValPheTrp---SerArgAspGly 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuArg---SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeu 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAsp 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnileLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGly 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 AlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIle---AlaAsnPro 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThraspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLys 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 IleproHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
                                                                                                                                                                                                                                                                                                    ----IleProLeuLeuAspTrpValGlyTrpAsn 61
                                                                                                                                                                                                                                                                                                                                                                                                           LysVallrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIle 99
                                                                                                                                                                                                                                                                                                                                                        62 AsnTrp-----GlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsn 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1941 ACGGCCACCGACCCTCCGGCAACACCCAGTGCGGGCAGCAGTGTCACCGTGGACTCGGTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyAlaThrTrpGlnAlaValProGlyAlaPro---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2289 GACGCAAC----GCCAACCCGATTGCCCAGGTCACCCCC----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-3 (1-740) x US-09-252-991A-13774 (1-4188)
                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13774
; LENGTH: 4188
                                                                                                          TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
                                                                                                                                                                         2.17e-08
203.00
32.38%
22.22%
5.03%
                                                                                                                                                                                                                                                                                                   47 AlaAsnGlyArgTrp
                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                        US-09-252-991A-13774
                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2181
                                                                                                                                                                                                                                                                                                                                                                                                             80
                                                                                                                                                                                                                                                                                                                                                                                                                                       2387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          217
                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 8
```

```
1821 CCGGGCAGCAGCGTGACCCTGACCGATGGCAACGGCAACCCGATCGGCCAGGTCACCGCC 1762
                                                                                                                                                                                        1761 GACGGCAGCGGCAACTGGAGCTTCACCCCGTCCACGCCGCTGGCGGATGGAACCGTGGTC 1702
                                                                                                                                                                                                                                                                                        1581 GCGGAACCGGGCAGCACGGTGATCCTCACCGACGGCAACGGCAATCCGATCGCCGAGGTC 1522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1521 ACCGCCGACGGCAGCGGCAACTGGACC-----TACACCCCCGT--CCACGCCGA 1476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1320 ACCGTGAGCCTGACCGACGGCAACGGCAACCCGATTGGCCAGGTCACGCCC---GACGGC 1264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1203 ACGGCCÁGCGACCÓGACCÓGCAATACCAGCÓCTCÓGGCCAGCACCACCGTGGACTCGGTG 1144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -- GCGCCGGCCGCCGGTGGTCAATCCGAGCAACGGAGTC 1105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1263 AGTGGCAACTGGAGCTTCACACCGGGACGCCGCTGGCCAACGGCACCGTGGTCAACGCC 1204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnPro 369
ThrGlyGlyProTyr-----AspGlySerSerGlyAspValTrpLysPheSerVal 277
                                                                                                                                                                                                                                                          -----SerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAsp 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----IleAspProPheAsnSerAspArgMetLeuTyrGlyThr------GlyAla 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------AspLeuThrLysTrpAspSer 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   415 GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LeulleSerProProSerGly-----AlaProLeulleSerAlaLeuGlyAspLeu 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1017 AGCGGCAACTGGAGCTTCACCCCGTCCACGCCGCTGGCGGATGGAACCGTGGTCAACGCC 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----ThrValAlaAla 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     307 ArgGlnHisProAsnThrIleMetVal------AlaThrGlnIleSerTrpTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              330 SerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     370 AsnProProValPro----SerProLysLeuGlyTrpMetAspGluAlaMetAla---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1436 CCGCCGGCAATAGCAGCCCCGGGCGCCAGCGTCACCGTGGAC-TCGCAAGCCCCGGCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 452 GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Accectacceaccceecceec---aaracceeceaccaeecaccacraccereeacec
                                                                                                                             278 Thr --- SerglyThrTrpThrArglleSerProValPro-----serGlyThrTrpThrArglleSerProValPro-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  517 GlnGlySerGluProGlyGlyValThrThrGlyGly----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ProAsp-----ThrilellePheArg-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrLeuTyrAlaThrAsn-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  512 GlyLysAsnTrp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1143 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              472
```

US-09-917-376-3 (1-740) x US-09-252-991A-13773 (1-4131)	Qy 48 AsnGlyArgTrp11eProLeuleuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65	Oy 66AsnGlyValValSexIleAlaAlaAspProlleAsnThr 78 ::	Qy 79 AsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp 93 :::	Oy 94 ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThr 112	Qy 113 ProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAla 132 .	Qy 133 ValAspProAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrp 152	Qy 153 ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr 172	Qy 173 TyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrp 192 :: 3642 CCGCTGGCCACCGGGATCAACGCAT3613	193 ValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePhe 194 ValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePhe 195 ValAlaPheAspLysSerSerSerCeGGGGGAATACCAGGGGGGGGGGGGGGGGGGGGGGGGGG	211ValGlyValAlaAspProAsnAsnProValPheTrpSer	224	3492 GCCGGGGCCACGGTGATCCTCACCGACGCAACGGCAACCGGTCGGCCAGGTCACCGCC	3432 GACGCCACGCCAACTGCGCTTTCACGCCCCCCCCCCCTTGCCCAATGCCACGCTGATC	Db 3372 AATGGGCTGGCCCAGGACGCCGCGGCAACAACAACAGCGCTCCCACCAGCGCCACCGTCGAC 3313	239 GlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThr 258	OY 259 SerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPhe 275	3252 GCCGAGGCTGCTCCTCACCGACGGCAACCGCAACCGGATCGGCCAGGTC	Qy 276 SerValThrSerGlyThrTrpThrArglleSerProValPro 289 :::	Qy 290SerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr 304	305 IleAspArgGlnHisProAsnThrIleMetValAla
544 GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsn 558		Gln	Qy 572 IleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSer 586 :::	PheGlnProvalAlaAla	607 LeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAsp	LeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSer	AlalleThrGlyValSerSerAlaValAsnValGlyPheGlyLysserAla	510 TTCACGCCGGGCCGCCCGCCGACGGCGTGATCGTCGCCACGGCCACGGCCGG 451 664 ProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla 683 [1]::::::	450 ACCGGCAATACCGGCCCGCCCCCCCCCGCGGGGGGGGGCGCCGCCGGCCC 392 684 TyrargSeraspaspCysGlyThrThrTrpValleulleabanaspaspGlnHisGlnTyr 703	391GGTGATCGAGCAGCAGCAGCAGCAGGCACGGCAC	349 GGAGGCCGGGGCCAAGGTGATCCTCACCGACGGCAA 308	RESULT 4 US-09-252-991A-13773/c Sequence 13773, Application US/09252991A Parent No. 6551795	GENERAL INFORMATION: APPLICANT: Marc J. Rubenfield et al. TITLE OF INVENTION: NUCLEIC ACID AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS	TILLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THEKAPEUTICS. FILE REFERENCE: 107196.136 CURRENT APPLICATION NUMBER: US/09/252,991A	CURKENT FILING DATE: 1939-02-18 PRIOR APPLICATION NUMBER: US 60/074,788 PRIOR FILING DATE: 1998-02-18	IOR FILING DATE: 1998-07-27 WHER OF SEQ ID NOS: 33142	SEQ ID NO 13773 LBNGTH: 4131 TYPE: DNA		2.92e-08 Length: 201.50 Matches: 31.62% Conservative:	Mismatches: Indels: Gaps:

ò	418	eHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerPr	438
qq	426053	TGACTTCGGCCCCTCGCCGGCGATCAACCTCAACATCGCCAA	426012
ò	438	OProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAs	458
qq	426011	GCCCGACGCTCCACCTGATCATCGTGGGCGCCCGGC	425969
ò	458	<pre>pValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerVa </pre>	478
qq	425968	AGCCGCCCGATCAGCCAT	425952
ò	478	IleIleValArg	498
qq	425951	TCCGATCATCGACTTGCGGCCAGCGCCCGGCTTCTTCAACGCCACCAC	425904
Vo	498	rGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyUsysAsnTrpPheGlnGl	518
Ор	425903	CGGCCCGTCGTCGGGCTTCCTCAACTGGGGTGCTG	425868
à	518	ySerGluProGlyGlyValThrThrGlyGlyThrValAlaAl	532
QQ	425867		425808
λ̈́o	532	aSerAlaAspGlySerArgPheValTrpAlaProGl	544
qq	425807	cadcateggaaaricdggcriccaaaactatgggtcgctgcagtcggcctgdggc	425753
λ̈	544	yAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTr	260
QC	425752	AAITTGGGCAACAGCAICTCGGGCAICTA	425724
ò	260	pAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLy	280
Dp	425723	: :	425673
ò	580	sThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVa	598
Dp	425672	CACCAACCTGGCTGGGTTGCAGAACGGCCCG	425639
λ̈	598	ThrPheGlnProValAlaAlaGlyLeu-ProSerSerGlyAlaValGlyValMetPhe-	617
Dp	425638	-ACGAGACGACCTTCAGCGTGGGCTTGGCCACCTCGGGTTCTGGAATCTGGGTAG	425583
Š	618		630
Ωp	425582	CGCAAACATCGGCAACTACAACCTGGGCGCCAACATCGGCGTCTACAACCTGGGCAG	425523
à	630		647
дg	425522	ĠĠĊĊĸĸĊĸĸĊĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸĸ	425463
à	647	laileThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerS	299
Dp	425462	caccaacaregeractec-aacaredsireseaacsreserecedes	425414
à	667	erTyrProAlaValPheValValGlyThrIleGly~~~GlyValThrGlyAlaTyrArgS	989
OD	425413	CTGACGGCGGCCATCGGCAACACGGCTTCGGCAACACGGAAACGGAAACA	425362
ò	989	erAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnT	902
qq	425361	TCGGCATCGGCAATACCGGCACCGGCAACATCGGCTTCGGCAACACGGAAACA	425302
ò	106	rpGlyGlnAla1leThrGlyAspHisAlaAsnLeuArgArgValTyrIle-GlyThrA	725
qq	425301	roddaroggereaccedcaccarcargaccedercedeccedaroggered	425242
ò	725	snGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly 740	
අු	425241	GCAACATCGGGCTATTCAACTCCGGCACCGGCAACATCGGCTTCGGCAACTCCGGC 4251	186
PESTI	Y		

σ

```
3743484 CAGATTGACGGCTTCAGCACCCCGCTCACGATCGACCGCATCCCGCTGAACCTCGGC 3743425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         313
                                                                                                                                                                                                                                                                                                                                    3744162
                 ò
                                          셤
                                                                        ઠે
                                                                                                 ď
                                                                                                                                 ð
                                                                                                                                                         g
                                                                                                                                                                                      ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ਨੇ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3744701 CCAATGTCGGTGCGCAGAATTTGGGTGCTGCCAACGCCGGTAGCGGGAATTTCGGTTTCG 3744642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3745039 GGCGGTGGCGGTCAACCGGATGGC-----GATGCGGGCGTTGGCGGTGTCGAATCT 3744989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3744988 GTTGGGGCAGAACGCCGCAGCGATCGCGGCCGTCGAGGCCGAGTACGAGTTGATGTGGGC 3744929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3744868 GCCGGCGTTCAGCCCACCGGCGCAGG-CGTTGGGGGGGTGGTGTCGTGCGCGTTCCTCAATG 3744810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3744749 ATTACAACGTCGGGTTGGGCAATGTCGGGATAT------TCAACCTGGGCGCAG 3744702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3744928 CGCCGATGTGGCGGCGATGGCCGGCTACCATTCCGGCGCGCGTCGGCTGCTGCCGCGCGCTT 3744869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3744809 CTCTATTTGCCGGACCCGCGAAGATGTTGAGGCTTAACGCGGGCTTGGGCAATGTCGGTA 3744750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3744581 TGGGCAATATTGGGTTGGGCAATGCGGGCAGCAGCAACTACGGCCTCGCAAACCTGGGTG 3744522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---SerSerAspGlnGlyAlaThr 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------Gly 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..... 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----AlaAlaValGlyMetTyr 88
                                                                                                                                                                                                                                                                                      at various positions throughout the sequence a, t, c or \boldsymbol{g}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------Irp---IleProLeu 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsp 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeu--- 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----TyrValArgThrAspileGly
                                       APPLICANT: FLEISCHANN, Robert D.
APPLICANT: FLEISCHANN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: PASER, Claire M.
APPLICANT: TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 2436-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --GlyMetTyrArgTrpAspAlaAlaAsnGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ThrAsnSerTrpAspProAsnAspGlyAlalleLeuArg-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             204
116
306
361
50
                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-3 (1-740) x US-09-103-840A-2 (1-4403765)
                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProlleAsnThrAsnLysValTrp----
                                                                                                                                                                                                                              TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at varid
OTHER INFORMATION: represent a, t, c
Sequence 2, Application US/09103840A Patent No. 6294328 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                             0.0564
184.00
32.65%
20.82%
4.56%
                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                     US-09-103-840A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
3744402 TCCGGCACCGGCAACATCGGGTTCTTCGGGGACCGGCAACTTCGGGGTATTCAAC 3744343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3743926 AACCTTCAGCCTCGGCCC---GCTGGTCGTTCCGGACATCGTGATCCCCGCTTTTGGTGG 3743870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3743869 CGGTACGGCCATACCTATCAGCGTCGGCCCCATCACCATCTCGCCCATCACCCTGTTCCC 3743810
3744521 TGGGCAACATCGGTTTTGCCAACACGGGTAGCAACAACATCGGGATCGGGTTGACCGGGG 3744462
                                                                                                                                            3744461 AC-AACCTGACTGGCATTGGGGGCCTGAATTCAGGAACCGGTAATCTGGGGTTGTTCAAC 3744403
                                                                                                                                                                                                                                                                                                                                                                                                                                         3744342 TCGGGCAGCTACAACACCGGTGTCGGTAATGCGGGGACGGCCAGTACCGGGTTGTTCAAC 3744283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3744223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3744103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3744102 TCG---GGCAACTTCAGCAACGGTGTGCTGTGGCGGGGTGACTACGAGGGCCTGTGGGGG 3744046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3743985 GG-CCCCATCACCGTTGCCGATCCAGATTTTGCCCACCATCCCGCTCAACATTTACCCA 3743927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3743690 TTAATATCGACACCAGAATTAACCAGTCATTCCCGGTGACCGTCAACTGGAGTACCCCGG 3743631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3743750 ACATTTCAGGAATCGAAATCAAAGATCTTGCCGGCAACGTCACCCTCCAATTAGGTAACC 3743691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GlyThrTyrIle 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             175 AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAla 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241 IleProHisLysGlyValPheAspPro-----ValAsnHisValLeuTyrIleAlaThr 258
                                                                    ---ArgGlyMetGlyGluArgLeuAlaValAsp 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AspThrileIlePheArgSerThrAspGlyGlyAlaThrTrpThrArglleTrpAspTrp 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||||
3744282 GTTGGTGGGTTCAACAGGGGTGTGGCCAACGTGGGTAGCTATAACACGGGGAGCTTCAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCGGCAACACCAACACCGGCAACTCGCCAACTCGGGCAATGTCAACACCGGCGCGTTCATC
                                                                                                                                                                                                                135 ProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLys---GlyLeuTrpArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -GlyGlnAlaSerLysThrIlePheVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              212 GlyValAlaAspProAsnAsnProValPheTrpSerArgAsp----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              226 -----GlyGlyAlaThrTrpGlnAlaValPro-----GlyAlaProThrGlyPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----TrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3743543 GGATCCGGGCTTCACCATTCCCGCTGCGCTGCCGCTG-ACGATCGACATAGACGGC
                                                                                                                                                                                                                                                                                                                                                                    -----ThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------ValProSerThrAspThrAlaAsn-----Asp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3743630 CAGTAACGATCTTCCCGAATGGCATCAGTÁTTCCCAACAÁTCCACTGGGGGTGCTGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThr---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360 Glu------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           259 SerAsnThrGlyGlyProTyrAspGlySerSerGlyAsp----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---SerGlyThrTrpThrArgIleSerPro----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              195 PheAspLysSerSerSerSerLeu-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           274 LysPheSerValThr-----
                                                                        GlyAsnMetProGly--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     344 ThrSer-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3743570 ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     170 Val-----
```

	RESULT 7 VE-90-328-352-1377 Sequence 1377, Application US/09328352 Patent No. 6562958 GENERAL INFORMATION: APPLICANT: Gary L. Breton et al. TITLE OF INVENTION: NUCLEIC ACID AMINO ACID SEQUENCES RELATING TO ACINETOBACTER TITLE OF INVENTION: BADMANNI FOR DIAGNOSTICS AND THERAPEUTICS	; FILE REFERENCE: GTC99-03PA ; CURRENT APPLICATION NUMBER: US/09/328,352 ; CURRENT FILING DATE: 1999-06-04 ; NUMBER OF SEQ ID NOS: 8252 ; SEQ ID NO 1377 ; LENGTH: 11679 ; TYPE: DNA ; ORGANISM: Acinetobacter baumannii	US-09-328-352-1377 Alignment Scores: 1e-05 Length: 11679 Pord. No.: 181.50 Marches: 204 Percent Similarity: 31.05\$ Conservative: 90 Best Local Similarity: 21.54\$ Mismatches: 309 Query Match: 4.50\$ Indels: 309	Gaps: 54)-328-352-1377 (1-11679) .alleGlyGlyGlyPheValAspGly ist:	Qy 221leValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAsp 37 iii	Qy 58 ValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAspPro 75		Qy 103 SerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 121 Db 1930 GGTGACGAGGCAATTGCTACAGATCCA
	Qy 406 ThrAsnAspLeuThrLysTrpAspSerGly	Qy 437 SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyClyPheThrHis 456 IIIIII :::::	477 SerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspPro 496 :::	516 eGln	37 58	GCCAAIGCCGGCACGGCA	604 aAladlyLeu	Qy 630 aAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSe 644 Db 3742633 CGGGACCGGGAACTTCGAGTTCTAACTCGGCAGCTTCAACACCGGCATCGCCAATGG 3742574 Qy 644 rTrpSerAlaileThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaPr 664 Db 3742573 CGGAACGGCAGTACTGGGCTTTCAATGCCGGTAATTCAATACCGGTGTGGCCAACCC 3742514

99

86

```
19060 ---CTGGATCTGGCGGGGGAAACCCGGTGGTGGAAAACGGTGGTCGATACG 19116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----GICCIGACGGICGAIACCICIATICCGACCACGCIGGCGCAAAICACCAGC 19470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19471 CAGACCACGCGGATACTACGCCGATTATTAGCGGGGTGATCACCGCCGCGCTTGCCAGT 19530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19195 CAG------GCCACCGACGACGACTACGCCGCTGCTGAACGGTGTACTTTCC 19239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18727 ACCACCTGGCAGGTGCGGGTGGTC----GATCTGGCGGGCAACGTT---GGCGCAACG 18777
                                                                                                                                                                                                                                                                                                                                                                                       18838 ATCAGCACCGACACGGGGAGTTCGGCAACTGACTTTATCACCAGCGACACCATGCTCACG 18897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18898 CTGACCGGTTCGCTGGGGGCGGGCTTGCCAGCGCGAAGTGGCGCAGATTAGCCTTGAT 18957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----CAGCGGCAGGACATTAAGCAGTTTG 19194
                                                                                                                                                                                               87 MetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGly 106
                                                                                                                                                                                                                                                                            107 AlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGly 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsn 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -- AspAsnileLeuTyr 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     143 Phe------GlyAlaProSerGlyLysGlyLeuTrpArgSerThrAsp 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             177 ProThrAspThrThrGly-----TyrGlnSerAspIleGlnGlyValValTrpVal 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      231 GlnAlaValProGlyAlaProThrGlyPheIleProHisLySGlyValPheAspProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThr---
                                   47 AlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsn
                                                                                                                 67 GlyvalvalSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGly
                                                                                                                                                                                                                                    18688 ------TGGAGTTACGTTGACGGACGACCCTCACCACGACGCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         194 AlaPheAspLysSerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----SerSerGlyAspVal-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----ValThrSerGlyThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         282 TrpThr-----ArglleSerProValProSerThrAspThrAlaAsnAspTyrPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :::|||
19360 TATACCTACAGCGCGAGTGGTGGTGTTTGGCGGGGAATATCACCTCCTCCAGTGATTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19117 ATTAACCCCACCGCCACACCAACGATTGTGTGTGTATACCGATGATGTCGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  251 AsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19018 CGCACGCTGACGACGACGCTACGTTTATCAGGTGCGGGTG--
18607 GGCGCTAATGAGTTCGCGCAAATCAGTACCGACAACGGCGCA
                                                                                                                                                                                                                                                                                                                                                        127 MetGlyGluArgLeuAlaValAspProAsnAsn-----
                                                                              ------ACCTGGGTGAAC--
                                                                                                                                        268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          599
                                                                                                                                                                                                                                                                                                                                                                                                                                        138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               273
                                                                                                                                                        18661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19168
                                                                            18649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                          셤
                                                                                                                                                    g
                                                                                                                                                                                               ઠ
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                               ठे
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                       ò
                                                                                                                                 1117 GGTGATACAGTAACTGCAACAGCAACTGACCTGCAGGCAATACATCGTTGCCAGGTACA 4176
                                                       1057 ACAGTGGTAGCAGGCACAGATGGTAGCTGGTCAGTACCAAACCCCAGGTAACCTGGTAGAT 4116
                694 ValLeuIle-----AsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThr 711
                                                                                              727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27 GlyAlaProGly11eLeuTyrValArgThrAsp11eGlyGlyMetTyrArgTrpAspAla 46
                                                                                                                                                                                                                                                                                                                                                                      Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Quarles & Brady
STREET: 1 South Pinckney Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
COMPUTER: WAR PC-DOS/MS-DOS
SOFTWARE: WAR PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLIASSIFICATION ATA:
PRIOR APPLICATION DATA:
                                                                                            GlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsn------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740) x US-09-453-702B-39 (1-25165)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: DNA (genomic)
DESCRIPTION: SEQ ID NO: 39:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Blattner, Frederick F
Burland, Valerie
                                                                                                                                                                                                                4177 ĠĠĊACAĠTĊTCTGCAĠĀĊĀTĊ 4197
                                                                                                                                                                                                                                                                                            Sequence 39, Application US/09453702B Patent No. 6365723 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3: (608) 251-5000
(608) 251-9166
                                                                                                                                                                         728 GlyllevalTyrGlyAspile 734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELECOMMUNICATION INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 39: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /.000337
170.50
29.22%
20.40%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 25165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: Madison
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: WI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
SEQUENCE
US-09-453-702B-39
                                                                                                                                                                                                                                                                       JS-09-453-702B-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
DB:
                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                               셤
                                                   셤
                                                                                                                                                                         à
```

0	Alignment Scores: 1.09 Length: 4403765 Pred. No.: 170.00 Matches: 171
GGGCAGTATGTTCAAGTGCTGATCAAGGCAAAACCTACACCTCTGAACCGGGGGGCGCGCGC	5/8

3792489 AACCCCGAGATGCTGGCAGCGGCGACCGATTTGGCCGGCATCCGGTCGGCGATCAGC 3792548 369 -ProAsnProProValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaileAs 388	/ 388 pProPheAsnSerAspArgMetLeuTyrGly	/ 399 -ThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGln11 418	418 eHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerPr 438 :::	438 oProSerGlyAlaProLeuIle	446 rAlaLeuGlyAspLeuGlyGlyPhe-ThrHisAlaAspValThrAlaValProSerT 465	465 3792908	/ 485 roserllelleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisV 505	for alAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyV 524	524 althrthrGlyThrValAlaAlaSerAlaAspGlySerA 538 ::	7 538 rgPheValTrp-AlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheG 557 11:::::	557 lyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgV 577	577 alAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyG 597	597 lyvalThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetP 617	617 heHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrH 637	(637 isSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnV 657	/ 657 alGlyPheGlyLyBSerAlaProGlySerSerTyrProAlaValPheValValGlyThrI 677	677 leGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeulleA 697 ::
# Conservative: 82 Db # Mismatches: 324 Indels: 184 Gaps: 32 Db	Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Strange Stra	PheLysLeuGlyGlyAsn:::			177	rpValAlaPhe GGCACCCGCCA) ρ, ι	spGlySer -		AspTyrPheGlyTyrSerGlyLeuThr	TOTAL TRANSPORT TO THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF) I A	345	379242	ITCGACCCATTGAGGGAGAACAGCCATGTGA-TCGCG 3/92488 QYTrpLeuThrPheGlyValGln 368 :::
Percent Similarity: 33.42% Conservative: Best Local Similarity: 22.59% Mismatches: Query Match: 3.15% Gaps: INS. 09.917.376.3 (1.740) y INS. 009.103.9400.2 (1.440.376.8)	Qy 82 TrpalaalavalGly Op. 3791642 TrpalacalGly Op. 3791642 TrpalacalGly Op. 3791642 TrpalacalGly Op. 3791642 TrpalacalGly	3791				Qy 178 ThraspThrThrGly Db 3791884ACCACAGG	Qy 196 AspLysSerSerser : : : D 3791938 TCGAACGGATCACC	3791998	3792058	249 ProValAsnHisVa	269 SerGlyAspValTr 3792148 GACGCATCAACCAC		305	317	337		Qy 359 aGluPro

**************************************	ě	
OY 097 SIMPEPARPALINITED AND TOTAL TOTAL TOTAL TOTAL AND THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF	à a	<pre>414 Atamspriohanmanriovatrneirpserarghapgis</pre>
Qy 717 euargargValTyrIleGlyThrAenGlyArgGly728 	& A	231GlnalaValProGlyAlaProThrGlyPhelle 3801236 CACGCAAGGGAAACCCAGCGCCACCGCGCGTTCC
Oy 729IleValTyrGlyAspileGlyGlyAlaProSerGly 740 ::: ::	& g	249 ProvalAsnHisValLeuTyrIleAlaThrSerAsnThi ::: ::: 3801287CATTCCGCGACAAGATCAAA
RESULT 10 US-09-103-840A-1 ; Sequence 1, Application US/09103840A . Datont No. 6204328	රු සි	269 SerGlyAspValTrpLysPheSerValThrSerGlyTh1 ::: 3801326 GACGCATCACACCTCAAGCGCAGCTACGGCTGGAACC
GENERAL INFORMATION: ; APPLICANT: FLEISCHMAN, Robert D. ; APPLICANT: PLEISCHMAN, Robert D. ; APPLICANT: PLEISCHMAN, Robert D.	රු වූ	286SerProValProSerThrAspThrAlaAsnAspTy3801386 CCGGCGCCCGAACCTGGTGCGGACACGCGCCGACTTGTGCGACACGCCGTCT
APPLICANT: VENTER, John C. ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM ; TITLE OF INVENTION: TUBERCULOSIS ; TITLE DRIBBERGERIADA: TUBERCULOSIS	6 6	305 IleAspArg-GlnHisProAsnThrIleMet3801431 TCGTCAAGATCAGCACCCTGGCAGCGTGACAGATCACCC
CURRENT ABPLICATION NUMBER: US/09/103,840A CURRENT FILING DATE: 1998-06-24 NUMBER OF SEQ ID NOS: 2	& A	317 rGlnIleSerTrpTrpProAspThrIleIlePheArgSe
; SOFTWARE: Patentin Ver. 2.1 ; SED TO NO 1 ; LENGTH: 411529 ; TYPE: DNA	& ଶ	337TrpThrArglleTrpAspTrpThrSer3801548 TTAGTGGCCGCCGCTTTGGGTCCACCGGGCCCTGCGG
; OKGANISM: Mycobacterium cuberculosis ; OTHER INFORMATION: H37Rv US-09-103-840A-1	yo e	346ArgSerLeuA) : :
1.35 Le 169.00 Ma 33.42% Co	8 8	
Mismatches: Indels: Gaps:	\$ 6	369 -ProAsnProProValProSerProLysLeuGlyTrpMe
US-09-917-376-3 (1-740) x US-09-103-840A-1 (1-4411529) Qy 82 TrpAlaAlaValGlvMetTvrThrAenSerTrpAspProAsnAspGlvAlaIleLeuArg 101	8 8	
3800820 TGGGCGGCGAAAGCCTGGCTACGCCAGGCCATCAACCATCTCGA	ብ <i>ኤ</i>	3801786 GGTGTCGCTCGCCATCTCGGCCCTGTTTGGCCAGCACGC 3801786 GGTGTYAGATATATATATATASPASPLeuThrb
Qy 102 SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 121 ::: Db 3800865 GAAGCTGATCGGACGCACCGAGCGGGTGGA 3800897	g 5	
Oy 122MetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsn 137	A 6	418 en.slleAraFromervalLysGlyLeuGluGluTnfa. :::: 3801906 GTATGCGGCCGCGAGAGCCACACCGTCGAGCAGATGGT
138 AspAsnileLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSer 157 ::::::::::::::::::::::::::::::::::::	à a	438 oProSerGlyAlaProLeulle 1
158 GlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnPro 3801014 TCGGCTACAGGCCCAGGTTGTGTGTGTGTGTGTGTGTGTG	රු සි	446 rAlaLeuGlyAspLeuGlyGlyPhe-ThrHisAlaAsp.
TrpValAlaPhe 195	& a	465 hrllePheThrSerProValPheThrThrGlyThrSerV
196 AspLysSerSerSerSerLeuGlyGlnAlaSerLysThrilePhevalGlyVal 213	& g	485 roSerllelleValArgAlaGlySerPheAspProSer6

```
CGCGCCCACCCGACCACGCCAC 3801490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ArgTyrValLeuAspIleSerAl 359
:::||||:::
||||
RGCCATGTCGTTTGTGA-TCGCG 3801666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MetAspGluAlaMetAlaIleAs 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GCCCAGGCCTATCAGGCGCTCAG 3801845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlavalAsnAspLeulleSerPr 438
:::||| :::||
GTGCTCAACGCGATCAACGCGCC 3801965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               p---ValThrAlaValProSerT 465
|||||||:::|||
GGCGGCAACGGCTTTACCCAGAC 3802085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGGTTGATCGCCAACGCCGGGC 3802145
                                                                                                                                                                            Ardeceaacce---carceaag 3801325
                                                                                                                                                                                                                                                  CCGCACCGAACTCACCGGCATCA 3801385
                                                                                                                                                                                                                                                                                                                    ------TCGCCCACAACC 3801430
                                                                                                                                                                                                                                                                                                                                                                                                                              SerThrAspGlyGlyAlaThr-- 336
|---||| || || :::
|GGGACTTTTTCAGGTCTTAGTAA 3801547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CCGGCATCCGCTCAGC 3801726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GCCTGACCTCCGGCGGCAACCT 3801905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGCGCCAACGGGACCGCGGAGAA 3802025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ...... 3802170
                                                                                                       CGAACACCGACGGG----- 3801286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GCGAAACACCAGACGTGATGCCG 3801607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuThrPheglyValGln---- 368
                                                                    leProHisLysGlyValPheAsp 248
                                                                                                                                        hrGlyGlyProTyrAspGlySer 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LysTrpAspSerGlyGlyGlnIl 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rValAspTyrAlaGluLeuAsnP 485
lygly---AlaThrTrp---- 230
                                                                                                                                                                                                                hrTrpThrArgile----- 285
                                                                                                                                                                                                                                                                                     yrPheGlyTyrSerGlyLeuThr 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----Se 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             rSerGlnProAsnAspArgHisV 505
```

1757

109

```
1532 ACCGATGGCGACGGC---CAGCCGGACACCACCGTGGTGGTCGCCCCGGCGGCAGGTTC 1476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GTGGTCAATGTGGTGGCCAGGAACGCCGCCGCAACAGCAGCCGGCG 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1592 GCGCCGGACGGCAGCAGCGTCACCGGCAACGCCGGGCGCCACCGTCGGTGTCGAC 1533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (379 derecedaetrecedaegegegegegenangechgenaegegegegeterengegen 1320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1319 ACGGCGGAAGCGGCGTGACCATCGTGATCACCGACGGCAACGGCAATCCGGATGGCCAG 1260
                                                                                                                                                                                                                                                                                                                                                               -----decedentreagerrecentaaceced 1713
                                                                                                                                                                                                                                                                                                                                                                                                                                   -----AsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGln 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       232 Ala-----ValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspPro 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---AlaThrTrpThrArg1leTrpAspTrpThrSerTyrProAsn 348
                                                                                                                                                                                                                                                                                                                              110 GlnIleThrProLeuProPheLysLeuGlyGlyAsnMet-------ProGly 124
                                                                                                                                                                                                           GAGGTCCCGCTGAACCCGCCGCTGACCAATGGCGAGACGGTGACGTGATCGTTACCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ThrGlnIleSerTrpTrp-----ProAspThrIleIlePheArgSerThrAspGlyGly
                                                                                                                                                                                                                                                            90 AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrp
                                                                                                                                                                                                                                                                                                                                                                                                   125 ArgGlyMetGlyGluArgLeuAlaVal-----AspPro----AspAspAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAla-----Phe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               196 AspLysSerSerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         250 ValAsnHis-----ValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr
                                                                                                                                                                                       72 AlaAlaAspProIleAsnThrAsn-----LysValTrpAlaAlaValGlyMetTyrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SerSerGlyAspVal-----TrpLysPheSer------ValThrSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 IleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAsp-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetVal----
                                2319
169
80
275
160
                                                                                                                                                       US-09-917-376-3 (1-740) x US-09-252-991A-13875 (1-2319)
                                                                  Conservative:
Mismatches:
Indels:
                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3.35e-05
164.00
36.56%
24.82%
                                                                                                                                                                                                                                                                                                                                                        1756 TG-GTCGCCCC----
                                                                                                                                                                                                                                                                                            1777 GCCAGCCGGACACCA
                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  216 Pro
                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  281
                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INCOMPATION:
APPLICANT: Marc J. Rubenfield et al
APPLICANT: Marc J. Rubenfield et al
TITLE OF INVENTION: NUCLEIC ACID AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: ARRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT PAPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR PILING DATE: 1999-02-18
PRIOR FILING DATE: 1999-07-27
NUMBER: OF SEQ ID NOS: 33142
SEQ ID NO 13875
LENGTH: 2319
                                                                                                                                                       3802402 -----TTGACCTCTTCGGTCGGCG 3802420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3802730 TCAACAACGCCGGTGGTGCCACCGGCGGCACCGGCGGTACCGGCGGGGCGCGGCGCTCTG 3802789
                                --GCCGGCGGCCTCGGCGGCAACGGCGGGTGGCTGTACGGCAACGCCGGGGCCCGGCGGCG
                                                                                                   3802229 TCGGGGGCGCGGGCACCGGAACCGGTGGTCACGGCGGGGCCGGCGGGGCCGGCGGGCCGGCGGG
                                                                                                                                                                                                                                        3802479
                                                                                                                                                                                                                                                                                                                                                                                                                                                    GCGGCGCCGGTGGCGGGGTACCTGGTCGGCGATGGCGGCGCCCGGCGGGACCGGCG 3802597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----ACCC 3802633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3802686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 717
                                                                                                                                                                                                                                                                                                                                                                                                                   637
                                                                lyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgV 577
                                                                                                                                                                                                                                                                                                                                                617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      637 isSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnV 657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         657 alGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrI 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             677 leGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeulleA 697
alAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySer---GluProGlyGlyV 524
                                                                                                                                                                                                                                                                                                                                             597 lyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetP
                                                                                                                                     rgPheValTrp-AlaProGlyAspProGlyGlnProValValTyrAlaValGly--PheG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 697 snAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnL
                                                                                                                                                                                                                                                                          alAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyG
                                                                                                                                                                                                                                                                                                                                                                               3802421 GCGCCGGCGGGGCCGGTGGCCACGCCGGGCTGTTCGGCGCCCGGCGGACCGGGGGACC
                                                                                                                                                                                                                                                                                                                                                                                                                 617 heHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3802790 GAGCGTGGCTGTACGGCAACGGCGGGGCCGCCCGGGGCCGGC 3802830
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----IleValTyrGlyAspileGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3802598 GGGCCGGCGGAAGAATTCCAGCGGTGGCGC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    euArgArgValTyrIleGlyThrAsnGlyArgGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 13875, Application US/09252991A
Patent No. 6551795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: DNA
CRGANISM: Pseudomonas aeruginosa
US-09-252-991A-13875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
US-09-252-991A-13875/c
505
                                                                                                                                                                                                                                                                          577
                                                                                                                                                                                                         557
                                                                                                                                     538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       729
                              3802171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3802687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                              셤
                                                                ò
                                                                                           셤
                                                                                                                                 ò
                                                                                                                                                                  g
                                                                                                                                                                                                    8
                                                                                                                                                                                                                               유
                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
```

1416

215

195

-----ACCGCCGAA 1380

1200

267

```
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-07-706-699-4
                                                             US-07-706-699-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . No. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            386
                                                                                                                                                                                                                                                        871
                                                                                                                                                                                                                                                                                                                      400
                                                                                                                                                                                                                                                                                                                                                                                   814
                                                                                                                                                                                                                                                                                                                                                                                                                                             413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1031 GGCAATCCGATCGGCCAGACCACCGCCGATGCCAACGGCAACTGGTCTTTCACGCCGTCC 972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---GTCACCGC 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YThrval-----AlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspPr 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---GlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnPr 602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        oVal-AlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProG 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   622 lyfysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyG 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :::||| :::||| :::||| :::|| :::|| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::|| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::|| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||| :::||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              642 lySerSerTrpSerAlalleThrGlyVal-----SerSerAlaValAsnValGlyP 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     659 heGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGly 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            462 CACCGTGTCAACGCCACGGCCACCGACGCCTCC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 166 AGGGCAGC---ACCACCGTCGATGGCGCGCGACCACGCCGACCGTCAACCTGAGCA
                                                                                                                          971 ACCCCGCTGCCGGACGGTA-------CCCGTGGTCAACGTGGTGGCCAGG
                                                                                                                                                                                           yValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAlaMetAl
                                                                                                                                                                                                                                                                                                                                                         ----AspleuThrLysTrpAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             nAspLeulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGlyGl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           453 yPheThrHisAlaAspValThr-----AlaValProSerThrIlePheThrSerPr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      oValPheThrThrGlyThrSerValAsp---TyrAlaGluLeuAsnProSerIleIleVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             lArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerTh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           oGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        349 ArgSerLeu-ArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPhe----Gl
                                                                                                                                                                                                                                                  GATGCCGCCGGCAACAGCAGTCCGCCGGCCGTAACCGTGGATGCCGTC-GCGCCGGC
                                                                                                                                                                                                                                                                                                                      a-----IleAspProPheAsnSerAspArgMetLeuTyrGlyThr-----Gl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ceccaececaacresaccrrcaccecesacececerrsececaaceseaesicaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAGTTCGCCGGCCAGCGTCACCGTGGACCCCGTGGCACCGGCCACGCCAGTGGTCAACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----AGCAACGCACCACGCTCAGCGGCACCGCCGAGCCGGCGCCCACCGTGACCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        rAspGlyGlyLysAsnTrp---PheGlnGlySerGluProGlyGlyValThrThrGlyGl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---TTGCCCAACGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  566 lProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      513 cáarádcagcadatradacricacrecaccacacác
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGCCGATGGCAACGGCAATCCCATCGGGCAG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ceccacectacceacccerceec----
                                                                                                                                                                                                                                                                                                                                                                                                                                          400 yAlaThrLeuTyrAlaThrAsn----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rAsn----
                                                                                                                                                                                                                                                        929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           428
                                                                                                                                                                                                                                                                                                                   386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                 ठे
                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
2670 CCTAAAGAGCCTGCGACAACACCTAAGAACCTAAGTGTTGTTAATGTAACTGAAACTACT 2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TyrThrAsnSerTrpAspProAsnAspGly-------AlaIleLeuArgSer 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----LeuLeuAspTrpValGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60 TrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProlle------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --AsnThrAsnLysvalTrpAlaAlaValGlyMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5741
177
98
276
394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-3 (1-740) x US-07-706-699-4 (1-5741)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ET AL-W2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ET AL-W2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
                               GENERAL INFORMATION:
APPLICANT: Schmid; Candussio; B ck
TITLE OF INVENTION: Maltopentaose
TITLE OF INVENTION: Producing Amylases
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Collard, Roe & Galgano, P.C.
STREET: 1077 No. 5204254thern Boulevard
COUNTRY: New York
                                                                                                                                                         E: Collard, Roe & Galgano, P.C. 1077 No. 5204254thern Boulevard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                       ZIP: 11576
COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/706,699
FILING DATE: 19910529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REGISTRATION NUMBER: 26,048
REFERENCE/DOCKET NUMBER: SCHMID
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-365-9802
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/POCKET NUMBER: 27,638
NAME: Bdward R. Freedman
REGISTRATION NEWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22,532
3ER: SCHMID
Sequence 4, Application US/07706699
Patent No. 5204254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           47 AlaAsnGlyArgTrpIlePro
                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: GR 40
FILING DATE: 31 MAY 1990
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Allison C. Collard
REGISTRATION NUMBER: 22,
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0.000168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Thomas M. Galgano REGISTRATION NUMBER: 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             163.00
29.10%
18.73%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4.048
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5741 bases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUCLEIC ACID
```

SEQUENCE (LENGTH: TYPE: STRANDE! TOPOLOG! MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (MOLECULE (M	Alignment Score Pred. No.: Score: Percent Similan	Best Local Sim Query Match: DB:	US-09-917-376-: QY 47 i	2562	2622	77	88	2730	103	2790	2850	125	2910	145	2967	163	3027	183	3051	203	3081	214	3141	216	3201	217	3261
- Sn	Alignm Pred. Score: Percen	Best Quer DB:	o-sn Yo	g &	g	8	8 &	qq	à	<u>අ</u>	ž 8	ò	qq	δ	qq	ò	qq	δ	qq	ò	qq	δ	qa	ò	Ор	ò	අ .
																											_
		638 FTTCTAGAGCAGTT 4943	leThrGlyValSer 652 ACGGGTGAGGAA 4997			IASPASPOYSELYINI 092 -GATGATGCAGTTGATCCT 5081													,								٠
eTyralaLeuSerAsnGlyThrPheTyrargSerThrAspGly	.ccaicccagascaassig. AspLeuTrp GTGAATCACTTTGGGGAGG	euTyrHisSer ::::: :GTTCCATTCAGGCTATA1	-ThrAenGlyGlySerSerTrpSerAlalleThrGlyValSer 	erAlaProGlySerSerTy	GGTAAAGGGTAT	ATGCTGATG					ducing Amylases	P.C.	5		¢							. AIW2	!	. AL-W2	W-,14		
ThrPheTyr	ProG1 TGGGG	LeualaalaSerSerGlyLeuTyrHisSer	ThrAenGlyGlySerSerTrpSer ::: GATGGTGTATTTCTGTTGACGATGGAGGAATAAATTATAAA	SeralaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPhe	GCAGGAATTACAGTAAAGTTTTCTAAA-	valvaleryiiiitteliyelyvaliiicilyalalyiatgeetabrabrabchyselyiiitiii TCAACATTGTGGGGGGTGTTAGATGCTGATGATGAGTTGATCCT	TrpValLeulleAsn 697		5011 13 Sequence 4, Application US/07998931	. N	ld; Candussio; BCk ON: Malopentaose Producing Amylases CES: 4	3 ADDRESS: Collard, Roe & Galgano,		41	E FORM:	IBM PC compatible	5.0	MBER: US/07/998,931	. 435 N D&T&:	MBER: GR 40 17 595.	NFORMATION:	UNBER: 22,532 WMBER: 22,532 KT NUMBER: SCHMID RT	638	ET NUMBER: SCHMID ET P Freedman	N. FICCAMAIN UMBER: 26,048 ET NIMBER: SCHMID ET	TELECOMMUNICATION INFORMATION: TELEPHONE: 516-365-9802	Q ID NO: 4:
	4/82 CCACAMACCAIT 617 PheHisAlaVal ::: 4827CATGATATG	629	639		4998 GCAGGAA		693 TrpVall		-931-4 e 4, Applica	GENERAL INFORMATION:	APPLICANT: SCHMIG; TITLE OF INVENTION: NUMBER OF SEQUENCES	ENG E. C	CITY: Roslyn STATE: New York	COUNTRY: USA	ER REAL	COMPUTER: IBM PC	SOFTWARE: WordPerfect	APPLICATION NUMBER: FILING DATE: 199212	δF	APPLICATION NUMBER:	H 5	26	NAME: Thomas M. Gal REGISTRATION NIMBER.	REFERENCE/DOCKET NUMBER: NAME: Edward D Presedmen	E E	COMMUNICATIO	ATION FOR SE
ò a ò i		ò a	ò 8			강 옵	රු සි	erm:m 13	US-07-998-931-4 ; Sequence 4, A	; Patent ; GENERAL	; TITLI; NUMB)	, CORR.		COUNT	COMPI	000	SOS	AP	I G	AP	INTER	, RE	NAN .	A RE	REC	TELE	; INFORM

```
||||
TGG-----TACGATGGCGTAAAGTGGCTTGATTCAGATCCTTTTGAAAGGGAA 2669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GCATCTAATGGTTAC-----GTAACAGGA 3050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |||:::
|TATGAGGTTCTGCGTGATGGTGTGGTT------ATT 3080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |||::: ::::::|||
GGCGAAACAACAGGACAACATTCATAGATACTGGATTAGATGCTGATAGGACCTATACG 3140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :::||||||
TATACGATTGTTGCTCTCGGAGATGGCGGGCAAAGTCTGATCCGAGCGAAGCGTTAGAA 3200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ::: | | | |
GATACCCCATATATGCATTATGGTCGAAGGTGGAGAGTGGACGATCGTTCCAGGAATT 3320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ||| ::: |||
|CCTAAAGAGCCTGCGACAACACCTAAGTGTTGTTAATGTAACTGAAACTACT 2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAGGATGTTGTTGCTTCAACTATTCGTACAACATTTACGGATGAAGACCTTAATCCAGAT 2849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTAACAACTCAAGAAAAACCAGAAGGAAATCTAGTAACAATATACTATAAAAAAGGCTTT 3260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     216
                                                                                                                                                                                                                                                                                                                                                                                TyrThrAsnSerTrpAspProAsnAspGly------AlalleLeuArgSer 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TyrGlnSerAsplieGlnGlyValValTrpValAlaPheAspLysSerSerSerLeu 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnAsnProvalPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGly--- 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerAspGlnGlyAla----- 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---ThrTrpGlnIleThrProLeuProPheLysLeuGlyGly-----AsnMetProGly 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGly 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----LeuLeuAspTrpValGly 59
                                                                                                                                                                                                                                                                                                                                                                                                                               rrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIle------
                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                        AlaAsnGlyArgTrpIlePro------
                                                                                                                                                                                                                                                                                                                    -3 (1-740) x US-07-998-931-4 (1-5741)
E CHARACTERISTICS:
H: 5741 bases
NUCLBIC ACID
DEBUESS: single
OGY: linear
E TYPE: DNA (genomic)
                                                                                                                                                               0.000168
163.00
29.10%
18.73%
                                                                                                                                                                                                                       arity:
milarity:
                                                                                                                                                                 res:
```

è	320	Qy 484 AsnProS	AsnProSerIleIl
5 1	7	Db 4281 GCTGGGT	GCTGGGTACAAAT
Q C	3321 AGAATGGAAGAATCAGAAATAGCAGGCTATAGTAAGTTAACCGTTGATATTCGGGAAGCA 3380	Oy 493	1
λ ₀	AlaProThrGlyPhelleProHisLySGlyValPheAspProValAsn :::	4341	ACGGATTCAGGCTT
a d	AGCAAGITUGAAGIAGCCTTTAATAATGGACGTGGGGCTTGGGATAGTGATCAAGAGAT	Qy 504 HisValA	HisValAlaPheSe
දු ද	252 H18ValleulYr 255 : ::	Db 4401 TCTAATA	 TCTAATAATTTCTC
a à	######################################	Qy 524 ValThrT	ValThrThrGlyGl
· 유	3501 ATTATTCCAGGTABAACCAGGAGCACCAATGGTAATAAAGTGACGATTTACTATCAA 3560	Db 4461 GTTGATC	GTTGATCCAGGAGG
ò	274	542	
qq	AATGGCTTTGATACGCCGTATG	4521	ACGCCAATCACAAT:
ò	SerProValProSerThrAspThrAl	.:: -	
DÞ	3621 CCAGGATTAAAAATGGAAGATTCAGAGTTTGCAAGTTATAGTAGGTTAACGCTT 3674	4561	A166C16
& 4	306 AspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThr 325	4641	CCAATGGACCTTAC
9 8	GAIA11UG1GAAGCIAA1CCIGCAGAAGIGGCI	Qy 564 GlnGlyVal	Val
S 6	ILELIER/NEARGSELINIASDOJYGIVALAINKIIDINKARGILEINDABDIIDINKSEL 34 	Db 4701 CAATATG	 CAATATGTAGATGT
§ &	TVYDYOA SARAKOS PILEBIO TANDA TANDA SARAT I PARATA PARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA SARATA	Qy 577 ValAsnP	ValAsnProLysTh
g q	GAAAATAATTATTTCTTCAATATTGGCGATAACACTTATATACCAGA 3		
ò	ThrPheGlvValGlnProAsnProProValProSer	Oy 597 GlyValT	GlyValThrPheGl
: 셤		Db 4782 CCACAAACCATT	ACCATT
ò	376ProLys 377	Qy 617 PheHisA	PheHisAlaValPro
: 셤		Db 4827CATG	-сатсататстс
ò	378 LeuGlyTrp 384	629	LeuAl
QQ	3912 GGTGGTACGTGGACAAATGCACCAGGAATAAAATGGATAAGGTCAGAAATAGCAGGTTAC 3971	4884	GATTACCAACTAAA
ò	MetAlaIleAspProPheAsnSerAspArg	639	
q	3972 AGTAAAATAACGCTTGATATTGGTCGCGCAGATCGAGTAGAAGTAGCCTTTAATGACGGT 4031	4944	GATGGTGTATTTCC
ò	399 ThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLygTrpAspSerGly 415	653	SerAlaValAsnVa
ДG	4032 CGTGGTGCA	4998	GCAGGAATTACAGT
ò	416	6/3	varvaigiyinriid
Бр	4062 AATTATCTCTTTGTAGTCGGTAACAATACTTATGAACCAGGAATTAACGGCGCACCTGGT 4121	5037	-TCAACATT
ò	430 ThrAlaValAsnAspLeulleSerProProSerGlyAlaProLeulleSerAlaLeuGly 449	Oy 693 TrpVall	TrpValLeulleAsi :::
QQ	4122 CAGGTGAAACATGGCGTGTTACCTGATGATGGAGAAGATCCGGGA 4166	PESITE 14	or serious
8 8	AspLeuGlyGlyPheThrHisAlaAspValThrAlaValPro	US-09-679-279-1 ; Sequence 1, Applica	ation US
Q Q	GATATTGAAGACCCTGATCATACCTCCCTTCAAGGCGGACTGATTTAACAGCAATAGCT	; Patent No. 6524841 ; GENERAL INFORMATION	ä
යි දි	464 SerThrilePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu 483 :::	; APPLICANT: McDaniel, Robe; APPLICANT: Volchegursky, ; TITLE OF INVENTION: Recor	iel, Rob egursky, N: Reco
! .;		; TITLE OF INVENTION	N: Gene

```
GGCATTAAGTGATGAACTGACAATTGAAAACCGCCGAGAAAACGGGT 4460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ||||:::::|||:::
GTGACGACCCACGTTCGCTCGGGGACAATTGGTCTATGCATGAAACA 4640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TAACAGATATTATTGACCCAGCTAACGCA-------4742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATAICGCATCTAATATGTTCCATTCAGGCTATATTTCTAGAGCAGTT 4943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :::|||||| :::::
CTGTTGACGATGGAGGAATAAATTATAAA----ACGCGTGAGGAA 4997
                                           rtrarccagarccrcaaarcccrcrracrcaarcaacrrar 4340
                                                                                                                                   || ::: ||| |- ::: ||| 3GGATATGCTTTGGTAAGAAGGTAACA 4520
                                                                                                                                                                                                                                                                                                                                                                                                                           -----GGCTCATCAGCTGGTACCACAATTAGCCAGATGGATATG 4781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----GCAATTGATACCATCCCAGAGCAAGGTGCAACA----- 4826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  raaagtititctaaa------sgiaaagggiat------ 5036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :: ||||||
IGTGGGGGGTGTTAGATGCT-----GATGATGCAGTTGATCCT 5081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gegegaaaaareeregaarcacrrregegaggaccagarraccr 4883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGly 596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMet 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          roglyLysGluGly----- 628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPhe 672
                                                                                                                                                                                                                                                                                        yThrValAlaAlaSerAlaAspGlySer----ArgPheValTrp 541
                                                                                                                                                                                                                                                                                                                                                                                               -----GlyAspProGlyGlnProValValTyrAla 553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------ SheglyAsnSerTrpAla----- 561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----AlaSer 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -------ProAlaAsnAlaGlnIleArgSerAspArg 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----ThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSer 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 leGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThr 692
                                                                                               ------SerPheAspProSerSerGlnProAsnAspArg 503
                                                                                                                                                                                                erThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGly 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   laAlaSerSerGlyLeuTyrHisSer------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hert
Y Anina
ombinant Megalomicin Biosynthetic
es and Uses Thereof
leValArgAlaGly---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S/09679279
```

```
INFORMATION: megAIII; SEQ ID NO: 15= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCATION: (22867)...(33555)
THER INFORMATION: megAII; SEQ ID NO: 14= translated amino acid sequence NAME/KEY: misc_faiture
LOCATION: (22957)...(24237)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCATION: (26230) ...(26733)
OTHER INFORMATION: megAII, KR3 (inactive)
NAME/KEY: misc feature
LOCATION: (26958) ...(27258)
OTHER INFORMATION: megAII, ACP3
NAME/KEY: misc feature
LOCATION: (27353) ...(28590)
OTHER INFORMATION: megAII, KS4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: misc feature
LOCATION: (37860)...(38120)
OTHER INFORMATION: megaIII, ACPS
NAME/KEY: misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION: (37068)...(37604)
OTHER INFORMATION: megAIII, KR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AME/KEY: misc feature
OCATION: (33052)...(33312)
THER INFORMATION: megAII, ACP4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATS
                                                                                                              LOCATION: (15427)...(16476)
OTHER INPORMATION: megAl, ATI
NAME/KEY: misc feature
LOCATION: (17155)...(17694)
OTHER INFORMATION: megAl, KRI
NAME/KEY: misc feature
LOCATION: (17947)...(18207)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION: (33780)...(35027)
OTHER INFORMATION: megAIII, KS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION: (21517)...(22053)
OTHER INFORMATION: megAl, KR2
MAME/KEY: misc. feature
LOCATION: (22318)...(22575)
OTHER INFORMATION: megAl, ACP2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YTHER INFORMATION: megali, AT4
AAME/KEY: misc. feature
(2953)...(3047)
YHER INFORMATION: megali, DH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OCATION: (24544)...(25581)
THER INFORMATION: megAII, AT3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION: megAII, KS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ION: (31396)...(32244)
INFORMATION: megAII, ER4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OCATION: (32257)...(32799)
THER INFORMATION: megali, KR4
                                                                                                                                                                                                                                                                                  CCATION: (17947)...(18207)
THER INFORMATION: megAI, ACPI
                                                                                                                                                                                                                                                                                                                               NAME/KEY: misc feature
LOCATION: (18268)...(19548)
OTHER INFORMATION: megal, KS2
                               LOCATION: (13849)...(15126)
OTHER INFORMATION: megAI, KS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCATION: (19876)...(20910)
OTHER INFORMATION: megAI, AT2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ION: (35385)...(36419)
INFORMATION: megAIII,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ... (29931)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ... (24237)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (33666) ... (43271)
     NAME/KEY: misc feature
LOCATION: (13849)...(1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: misc_feature
COCATION: (21517)...(22
                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: misc feature LOCATION: (19876)...(2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AME/KEY: misc feature OCATION: (24544)...(2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AME/KEY: misc_feature
OCATION: (26230)...(2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AME/KEY: misc feature OCATION: (28897)...(2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: misc feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: misc feature
                                                                                          NAME/KEY: misc feature
LOCATION: (15427)...(1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AME/KEY: misc feature OCATION: (31396)...(3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AME/KEY: misc feature OCATION: (32257)...(3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: misc feature
LOCATION: (35385)...(3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           megDIV, TDP-4-keto-6-deoxyglucose 3,5-epimerase (eryBVII, dnmU hd
TDP-4-keto-6-deoxyhexose 3,5-epimerase;
SEQ ID NO: 8= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (3462)...(4634)
JTHER INFORMATION: megG(megY), mycarosyl acyltransferase, mycarose O-acyltransferase
JTHER INFORMATION: SEQ ID NO: 5= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ...(6595)

W: megDIII, daunosaminyl-N,N-dimethyltransferase (eryCVI homolog);

W: SEQ ID NO: 7= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: (8228)...(9220)

DTHER INFORMATION: megBII-1(megDVII), TDP-4-keto-L-6-deoxy-hexose 2,3-reductase,

DTHER INFORMATION: SEQ ID NO: 10= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  megBV, mycarosyl transferase, mycarose glycosyltransferase;
SEQ ID NO: 11= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: (7220)...(8206)
7THER INFORMATION: megDV, TDP-hexose 4-ketoreductase (eryBIV, dnmV homolog),
7THER INFORMATION: TDP-4-keto-6-deoxyhexose 4-ketoreductase;
7THER INFORMATION: SEQ ID NO NO: 9= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: CDS
LOCATION: (1)...(144)
OTHER INFORMATION: megBVI (megT), TDP-4-Keto-6-deoxyglucose-2,3-dehydratase;
OTHER INFORMATION: SEQ ID NO: 2= translated amino acid sequence
NAME/KEY: CDS
LOCATION: (208)...(2061)
OTHER INFORMATION: megDVI, TDP-4-Keto-6-deoxyglucose 3,4-isomerase;
OTHER INFORMATION: TDP-4-Keto-6-deoxyhexose 3,4-isomerase;
OTHER INFORMATION: SEQ ID NO: 3= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             megDII, deoxysugar transaminase (eryCI, DnrJ homolog), TDP-3-keto-6-deoxyhexose 3-aminotransaminase, SEQ ID NO: 6= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: (12181)...(22821)
OTHER INFORMATION: megAI; SEQ ID NO: 13= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                megDI, rhodosaminyl transferase (eryCIII homolog),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION: (10483)...(11424)
OTHER INFORMATION: megBIV, TDP-hexose 4-ketoreductase,
OTHER INFORMATION: TDP-4-keto-6-deoxyhexose 4-ketoreductase;
OTHER INFORMATION: SEQ ID NO: 12= translated amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TDP-megosamine glycosyltransferase;
SEQ ID NO: 4= translated amino acid sequence
FILE REPERENCE: 300622004700.
CURRENT APPLICATION NUMBER: US/09/679,279
CURRENT FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/158,305
PRIOR APPLICATION NUMBER: US 60/190,024
PRIOR APPLICATION NUMBER: US 60/190,024
PRIOR APPLICATION NUMBER: US 60/190,024
NUMBER OF SEQ ID NOS: 34
NUMBER OF SEQ ID NOS: 34
                                                                                                                                                                                                                                                                                                        TYPE: DNA
ORGANISM: Micromonospora megalomicea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: misc feature
LOCATION: (13576)...(13791)
OTHER INFORMATION: megAI, ACP-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: (12505)...(13470)
OTHER INFORMATION: megAI, AT-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ... (10479)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (2072)...(3382)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (4651) ... (5775)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (6592)...(7197)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: misc feature LOCATION: (12505)...(1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: CDS
LOCATION: (4651)...(
OTHER INFORMATION: m
OTHER INFORMATION: T
OTHER INFORMATION: S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: CDS
LOCATION: (6592)...('OTHER INFORMATION: m'OTHER INFORMATION: TOTHER INFORMATION: S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: CDS
LOCATION: (9226)...(
OTHER INFORMATION: m
OTHER INFORMATION: S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: (5822)...
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OTHER INFORMATION:
                                                                                                                                                                                                                                                                                  LENGTH: 47981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFO
                                                                                                                                                                                                                                                  SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IAME/KEY:
```

Qy 120 yAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsn	162 rGlnMetThrAsnPheProAspValGlyThrTytleAla-AsnProThrAspT 17289 GCTACTCGCGAACTCGCCGACCTCGGGCCACCGCTCCGCGCGCG	1513 Shi svalleuTyrIlealaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyA 271 17543
LOCATION: (38187)(39470) OTHER INFORMATION: megAIII, KS6 NAME/KEY: misc feature LOCATION: (39795)(40811) OTHER INFORMATION: megAIII, AT6 NAME/KEY: misc feature LOCATION: (41406)(41936) OTHER INFORMATION: megAIII, KR6 NAME/KEY: misc feature COCATION: (42268)(42425) OTHER INFORMATION: megAIII, ACP6 NAME/KEY: misc feature LOCATION: (42585)(42371) COTHER INFORMATION: megAIII, TE NAME/KEY: CDS	CCATION: (43268)(44344) OTHER INFORMATION: megCII, TDP-4-keto-6-deoxyglucose 3,4-isomerase; OTHER INFORMATION: megCIII, TDP-4-keto-6-deoxyglucose 3,4-isomerase; OTHER INFORMATION: megCIII, TDP-4-keto-6-deoxyglucose; COCATION: (44355)(45523) OTHER INFORMATION: megCIII, desosaminyl transferase, desosamine glycosyltransferase; OTHER INFORMATION: SEQ ID NO: 17= translated amino acid sequence NAME/KEY: CDS COCATION: (45620)(46591) OTHER INFORMATION: megBII-2 (megBII), TDP-4-keto-6-deoxy-L-glucose 2,3 dehydratase; OTHER INFORMATION: SEQ ID NO: 18= translated amino acid sequence NAME/KEY: CDS LOCATION: (46660)(47403) COTHER INFORMATION: megH, TEII; SEQ ID NO: 19= translated amino acid sequence NAME/KEY: CDS LOCATION: (47411)(47980) COTHER INFORMATION: megF, C-6 hydroxylase; SEQ ID NO: 20= translated amino acid seque	Alignment Scores: Description

	S		
Db 19044 GGTGGGTTCGGCGGTGAATCAGGATGGGTAATGGGTTGGCGGCCGTCGGC 19100 Qy 636 TyrHisSerThrAenGlyGlySerSerTrpSerAlalleThrGlyValSerSerAlaVal 655 Db 19101	Oy 734 IleGlyGlyAlaProSerGly 740 Db 19293 TGTGGGTCATGTGCAGGCGC 19313 RESULT 15 US-09-255-991A-5730 ; Sequence 5730, Application US/09252991A ; Patent No. 6551795 ; GENERAL INFORMATION: APPLICANT: Marc J. Rubenfield et al. TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS TITLE OF INVENTION: AGROUNDSA FOR DIAGNOSTICS AND THERAPEUTICS FILE REFERENCE: 107196.136 ; CURRENT FILING DATE: 1999-02-18 ; PRIOR APPLICATION NUMBER: US 60/074,788 ; PRIOR PILING DATE: 1999-02-18 ; PRIOR PILING DATE: 1999-07-18 ; RIOR FILING DATE: 1999-07-18 ; ROWENT FILING DATE: 1999-07-27 ; NUMBER OF SEQ ID NOS: 33142	TYPE: DNA 1	
CCT 18064 rcG 398 TrG 18124 413 rcG 18184 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 418 ini 41	TCCGGGC 18355 PheThrs 469 ATGGTCT 18415 IlelleV 489	ThrolyG 528	
	18296 GCGGCTGCGGGGGGGGGCTCCCGGAGCAGCTGGGGAGTTGATCGTCTCCGGGC 1835S 449 lyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrS 469 18356 GGGACACCGCTCGGGGGCACCCGGGGACTCGGCGGAGTTGATGGTT 18415 469 erProValPheThrThfGlyThrSerValAspTyrAlaGluLeuAsnProSerIle11eV 489 18416 CCGACACGGGGGCGCCCCGTACCGCTTCGGCACTTCATGCCGGG		

```
2947 -----CTGTTCTACTTCGGCCCCGATAGCGCTGCGAGCCTGGCGATCTGCAACC 2999
                                                              -----ATCGTCGGCTCAATAG 3044
                                                                                                                                 --- 3075
                                                                                                                                                                                        3076 -------GGGAACCGCCGGCCGGGAACCTGGTATGAAGGTGCGGGTCCCGTCTG 3122
                                                                                                                                                                                                                                                                                               693 pValLeuIleAsnAspAspGlnHisGlnTyrGlyAsn---TrpGlyGlnAlaile---- 710
                                                                                                                                                                                                                                                                          711 ------HisAlaAsnLe 717
                                                                                                rAlavalAsnvalGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVa 673
                                                                                                                                                           673 lValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTr 693
                                           633 rGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaileThrGlyValSerSe 653
                                                                                                                     3000 GGCGCTTTCTATGCGTTGACAGAGAT---
                                                                                                                                                                                                                                                                                                                                                                                                                      3303 ¢GGCCCCGGT 3312
                                                                                                                                                                                                                                                                                                                                                                                          737 aProSerGly 740
                                                                                                     653
                                                                                                   ઠે
                                                                                                                                                                                    셤
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                     쉱
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                   ò
```

Search completed: May 12, 2004, 15:15:19 Job time : 12944.8 secs

Run on:

```
Sequence 1845, Ap
Sequence 1, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 2561, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 379, App
Sequence 37792, A
Sequence 14432, A
Sequence 14, Appl
Sequence 35048, A
Sequence 3833, Appl
Sequence 8383, Appl
Sequence 11, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 379, App
Sequence 97984, A
Sequence 97984, A
Sequence 97984, A
Sequence 9717, Ap
Sequence 21, Appl
Sequence 21, Appl
Sequence 21880, A
Sequence 1871, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 206, Applisequence 4, Applisequence 22, Applisequence 3375, Apsequence 42336, Apsequence 3, Applisequence 3, Applisequence 3, Appli
                                                                                                                                                                                                                                                                                                                                            Sequence 11871, A
Sequence 28185, A
Sequence 37642, A
Sequence 26288, A
                                                                                                                                                                                                                                                          Sequence 1, Appli
Sequence 30151, A
Sequence 3, Appli
Sequence 11269, A
                                                                                                                                                                     Sequence 21,
Sequence 13,
Sequence 11,
Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: VINZANT, TODD B.
APPLICANT: WINDER, MICHAEL B.
APPLICANT: WINDER, MICHAEL B.
APPLICANT: WINDER, WICHAEL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.4US01
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
US-09-917-376-2
US-10-155-400-2
US-10-155-400-2
US-10-156-761-1845
US-10-156-761-1845
US-10-156-761-1845
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-156-761-18
US-10-166-18
US-10-166-18
US-10-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: DNA ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/09917376
Publication No. US20040038334A1
GENERAL INFORMATION:
      3360
69350
5451
2110
2110
1839
3879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
NAME/KEY: modified_base
LOCATION: (2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-2
                                                                                                                                                                                                                1117
1114
1114
878.5
238
238
199.5
195.1
                                                                                                                                                                                                                                                                                                                                                                                      188.5
184
                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
179.5
176
170.5
170.5
170.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168
167.5
167.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          166
165.5
165
164.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161.5
161.5
161
159.5
159.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-MODEL=frame+ p2n.model -DEV=x1h
-Q=/Cgn2_1/USPTO_spool/US09917376/runat_11052004_114537_28336/app_query.fasta_1.1429
-Q=/Cgn2_1/USPTO_spool/US09917376/runat_11052004_114537_Ep2n.rnpb -MINNATCH=0.1
-DB=Published Applications NA -OFWT=fastap -SUFFTX=p2n.rnpb -MINNATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGEN=200 -THR SCORE=pct -THR MAX=100
-MAXIEN=2000000000 -USER=US09917376_GCGN 1 1 809_Grunat_11052004_114537_28336
-NCPU=5 -ICPU=3 -NO MAAP -LARGEQUERY NEG SCORES=0 -WART -DSPENDGCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                           May 12, 2004, 09:26:52 ; Search time 1136.23 Seconds (without alignments) 2950.286 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                            1 ATTQPYTWSNVAIGGGGFVD......YIGTNGRGIVYGDIGGAPSG 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Published Applications NA:*

| cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/PPT_NEW PUB.seq:*
| cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq:*
| cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
| cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
        GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                         OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                   2941586 segs, 2264995651 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                   BLOSUM62
Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length DB
                                                                                                                                                                                                                US-09-917-376-3
4036
                                                                                                                                                                                                                                        Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                       Scoring table:
```

Jatabase :

Searched:

Score

Result

; OTHER INFORMATION: a, c, t, g, other or unknown US-09-917-376-2	01 8	rGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 320
Alignment Scores: Pred, No: Score: Score: A 036.00 Matches: Percent Similarity: Best Local Similarity: 100.00\$ Mismatches: Query Match: 100.00\$ Indels: Gass DB:	321 321 1099 341	propertical and the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the c
x US-09-917-376-2 (1-2869)	1159	٦
Qy 1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAsp 20 	361 P	е н
Qy 21 Gly1leValPheAsnGluGlyAlaProGly1leLeuTyrValArgThrAspIleGlyGly 40		MetAspGlualaMetAlaIleAspProPheAsnSerAspargMetLeuTyrGlyThrGly 400
YTrp 6	Oy 401 Al	AlaThrLeutyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlyGlnIleHisIle 420
ASDASDTYDGIYTYASDGIYVAIVAISETIIEAIAAIAASDFOIIEASDTYS	1399	MetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer 4
Oy 81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAla11eLeu 100 	Qy 441 GI 	GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 460
ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 12	2y 461 Al	AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 480
AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnile 1	Qy 481 Al 	AlaGluLeuAsnProSexIleIleValArgAlaGlySerPheAspProSerSexGlnPro 500
LeuTyrPheGlyalaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr	Qy 501 As 	AsnaspargdisvalalapheSerThraspGlyGlyLygAsnTrpPheGlnGlySerGlu 520
	Qy 521 Pr 	roGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 540
ThrdlyTyrGlnSerAspileGlnGlyValValTrpValAlaPheAspiysSerSerSer 20	Oy 541 Tr 	rpalaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrp 560
SerLeuGlyGlnAlaSerLySThrIlePheValGlyValAlaAspProAsnAsnProVal 22	561	lasnProLys 5 gaarccaaag 1
C4 00	Oy 581 Th	ThrPheTyralaLeuSerAsnGlyThrPheTyrargSerThrAspGlyGlyValThrPhe 600
241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 26	Oy 601 GL 	GInProvalAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal 620
261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 28	Oy 621 Pr.	roGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn 640
OY 281 ThrIndThrArgileSerProValProSerThrAsapthrAlaAshAsivrPheGlyTyr 300	41	GlyGlySerSerTrpSerAla11eThrGlyValSerSerAlaValAsnValGlyPheGly 660
Db 979 ACATGGACGCGAATCAGCCCGGTACCTTCGACGACACGGCCAACGACTACTTGGTTAC 1038	661	089

```
ThrTrpThrhrg1leSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGCGGCCTCACTATCGACCGCCAGCACCCGAACACGATAATGGTGGCAACCCAGATATCG 1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          480
                                                                                                                                                                                                                                                                                                                                                                                                                       320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               460
                                                                                                                                                                                                                                                                                                                                 280
                                                                 160
                                                                                       618
                                                                                                           180
                                                                                                                                678
                                                                                                                                                      200
                                                                                                                                                                           738
                                                                                                                                                                                                 220
                                                                                                                                                                                                                     798
                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                858
                                                                                                                                                                                                                                                                                      260
                                                                                                                                                                                                                                                                                                           918
                                                                                                                                                                                                                                                                                                                                                     978
                                                                                                                                                                                                                                                                                                                              CCTTGGCTGACCTTCGGCGTACAGCCGAATCCTCCGGTACCCGAAGCTCGGAAGCTCGGCTGG
                                                                ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLy8SerSerSer
                                                                                                                                                                                                                                                                                      IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn
                                                                                                                                                                                                                                                                                                  TrpTrpProAspThr11e11ePheArgSerThrAspGlyGlyAlaThrTrpThrArg11e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     regrescessacacearancerressascaceaceaceareceacereseare
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGCGCCCGCTCATCAGCGCTCTCGGAGACCTCGGCGGCTTCACCCACGCCGACGTTACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr
  CGCTCGTCTGATCAGGCGCAACGTGGCAAATAACGCCCCTGCCGTTCAAGCTTGGCGGC
                     ASIMETProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsinAsinAsinEle
                                                                                                            TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr
                                                                                                                       ACCGGCTATCAGAGCGATATTCAAGGCGTCGTCTCGGCTTTTCGACAAGTCTTCGTCA
                                                                                                                                                                                               SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal
                                                                                                                                                                                                            TCGCTCGGGCAAGCGGGTAAGACCATTTTTGTGGGCGTGGCGGATCCCAATAATCCGGTC
                                                                                                                                                                                                                                           PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe
                                                                                                                                                                                                                                                           SerGlyLeuThr11eAspArgClnHisProAsnThr11eMetValAlaThrGlnI1eSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer
                                                                                                                                                                                                                                                                 199
                                                                                                                                                                                                                                                                                                                                  261
                                                                                                                                                                                                                                                                                                                                                      919
                                                                                                                                                                                                                                                                                                                                                                                                                                            1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1399
   439
                                             499
                                                                                                                                  619
                                                                                                                                                                                                                        739
                                                                                                                                                                                                                                                                                                                                                                                                 979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            461
                                                                  141
                                                                                        559
                                                                                                             161
                                                                                                                                                       181
                                                                                                                                                                           619
                                                                                                                                                                                                 201
                                                                                                                                                                                                                                             221
                                                                                                                                                                                                                                                                                        241
                                                                                                                                                                                                                                                                                                            859
                                                                                                                                                                                                                                                                                                                                                                             281
                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 441
                       121
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   & ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                  ద
                                                                                                           g &
                                                                                                                                                                        g
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        જે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                      8 8
                                                           ð
                                                                                                                                               ∂
                                                                                                                                                                                                8
                                                                                                                                                                                                                                     8
             2178
                                                        2238
                                                                                                 2298
                                                                                                                                             2358
                                                                                                                      740
                                                                              720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                           HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgVal
                                                                                                                      81 ValtrpalaalaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAla11eLeu

        ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly

                                              ACGGGGGCGTACCGCTCCGACGACTGTGGGACGACCTGGGTACTGATCAATGACGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              139 GCGACGACTCAGCCGTACACCTGGAGCAACGTGGCGATCGGGGGGGCGGCGGCGGCTTTGTCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGGATCGTCTTCAATGAAGGTGCACCGGGAATTCTGTACGTGCGGACGACATCGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MetTyrArgTrpAspAlaAlaASnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AACAATTGGGGGTACAACGGCGTCGTCGTCGCATTGCGGCAGACCCGATCAATACTAACAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GTATGGCCCGCCGTCGGAATGTACACCAACAGCTGGGACCCAAACGACGGAGCGATTCTC
ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Argraticaticacidececeaececeaececesecerecerecerecearidecerecearice
                                                                                                                                                                                                                                                                  FROM ACIDOTHERMUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2869
740
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                              RESULT 2

US-10-155-400-2

Squence 2, Application US/10155400

Publication No. US20030108988A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: HIMMEL, MICHAEL E.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: NREL 0136A

CURRENT FILING DATE: 0202-10-22

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PARENTING PATE: 2.1

SEQ ID NO 2

LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                        or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-3 (1-740) x US-10-155-400-2 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Acidothermus cellulolyticus FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                        other
                                                                                                                                                                                                                                                                                                                                                                                                                         ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4036.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                        ť,
                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (2869)
OTHER INFORMATION: a, C,
                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: modified_base
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
Score:
Percent Similarity;
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                         ; OTHER INFORM
US-10-155-400-2
                                                       2179
                                                                                                  2239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       319
             2119
                                                                              701
                                                                                                                                                                                                                                                                                                                                                                   TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      101
                                   681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                            g
                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요 장
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
             合
                                                 qq
                               8
                                                                    8
                                                                                                                      ò
```

g

⋧

ò

g ò g ò g à g à qq

```
AspileGinGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAla 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   266 AspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIle 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Grecresaceceses de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de constructor de co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 115 TACAGCTGGAAGAAGGCCCGCGTCGACGGCGGCGGCGTCGTCCCCGGCATCGTCTTCAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACGCGGACGGTGTACGTCGCGGACAAGGACAACTCCGTCTATCGCTCCACGGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGCACGTACACAGAACAGGTGGGACCCGGGCAACGGTGCCGTGCTCCAGGTCCGGCGACCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |||||||| GCGCGAGCTGGCCTTCAAGCTGGGGAACATGCCGGGCCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCCAGGGGGCAAGGGGGGTGGACGGACGGACGGGGGCCTCCTGGTCGCAGGGTCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 295 ACGGGTGTCGTGAGCCTCGCCTCCGACTCCGTCGACCACAACAAGGTGTACGCGGCCGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6 TyrThrTrpSerAsnValAlalleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLysGly
                                                                                                                                                                                                                                                                                                                                                         2646
438
111
174
19
                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-3 (1-740) x US-10-156-761-1845 (1-2646)
                                                                                                                                   avermitilis
                                                                                                                                                                                                                                                                                                                                               2.58e-229
2429.50
73.99%
59.03%
60.20%
         NOS: 15109
, NUMBER OF SEQ ID NOS: 1510;
SEQ ID NO 1845
LENGTH: 2646
TYPE: DNA
TYPE: DNA
CRGANIEM: Streptomyces av
                                                                                                                                                                                                                            ; LOCATION: (1)..(2646)
US-10-156-761-1845
                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                  NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 206
                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <u>ک</u> ۾
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P & P &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2118
                                                                                                                                                                                                                            1698
                                                                                                                                                                                                                                                                                                                                                         1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2358
                                                                                                                                                                                                                                                                                          540
                                                                                                                                                                                                                                                                                                                                                                                                                       560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               740
                                                                     AACGACAGGCACGTCGCGTTCTCGACACAGACGAGCAACTAGTTCCAAGGCAGCGAAA
                                                                                                                                                                                                                                                                                                                           CCTGGCGGGGTGACGACGGCGCACCGTCGCCGCATCGGCCCACGACTCTCGTTTCGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GCTGCTTCGCAAGGTGTTCCCGCCAATGCCCAGATCCGCTCAGACCGGGTGAATCCAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACGGGGCGTACCGCTCCGACTGTGGGACCTGGGGTACTGATCAATGACCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TACATAGGCACGAACGGCCGTGGAATTGTATACGGGGACATTGGTGGCGCCGTCCGGA
                                AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro
                                                                                                                                                       AsnAspArgHisValAlaPheSerThrAspGlyGlyuyaAsnTrpPheGlnGlySerGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACTITCIAIGE CONTACTATION ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTITION OF THE CONTACT ACTIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GInProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGCGGCAGCAGTTGGTCTGCAATCACCGGGGTATCCTCCGCGGTGAACGTGGGGTTTGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CACCAATACGGAAATTGGGGACAAGCAATCACCGGTGACCACGCGAATTTACGGCGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TyrileGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly
                                                                                                                                                                                                                                                                                          ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal
                                                                                                                                                                                                                                                                                                                                                                                                                    TrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ccriscaladadadeccarcricricricricricricricricalicalical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 3
US-10-156-761-1845
i Sequence 1845, Application US/10156761
j Publication No. US20030119018A1
i GENERAL INFORMATION:
j APPLICANT: OWURA, SATOSHI
j APPLICANT: IKEDA, HARUO
j APPLICANT: IKEDA, HARUO
j APPLICANT: IKEDA, HAROSHI
j APPLICANT: SHIRA, TADAYOSHI
j APPLICANT: SHRAKI, YOSHIVUKI
j APPLICANT: SHRAKI, YOSHIVUKI
j APPLICANT: SHRAKI, YOSHIVUKI
j APPLICANT: HATTORI, MASAHIRA
j TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
j TILE REFERENCE: 249-262
j CURRENT FILING DATE: 2002-05-29
j PRIOR APPLICATION NUMBER: US/10/156,761
j CURRENT FILING DATE: 2001-05-30
j PRIOR APPLICATION NUMBER: JP 2001-272697
j PRIOR FILING DATE: 2001-08-02
                                                                                                                                                                                                                            1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2179
                                                                                               1579
                                                                                                                                                             501
                                                                                                                                                                                                                                                                                          521
                                                                                                                                                                                                                                                                                                                                                  6691
                                                                                                                                                                                                                                                                                                                                                                                                                    541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           721
```

d

à

g

ò

엄

ò

g

ò

g

ð

g

à

В

ò

105

354

85

414

174

25

234

45

294

65

125

474

145

534

165

594

185

654

714 225 774 245

6 6 6 6 6

ò

```
2277508 TACAGCTGGAAGAACGCCCGCGTCGACGGCGGCTTCGTCCCCGGCATCGTCTTCAAC 2277567
                                                                                                                                                                                                                                                               2238
                                                                                                                                                                                               2178
 2239 CGCGTGTATGTGTCGACGAACGGCGCGGGATCGTCTACGGCGACACCGCGGGCTCTTCG 2298
                                                                                                                                                                 698
                                                                                                                                                                                                                                                                                            ArgValTyrIleGlyThrAsnGlyArgGly1leValTyrGlyAspIleGlyGlyAlaPro 738
                                                                                                                                                                                  AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnLeuArg
                                                                                                                                                                                                                                                ||| ||||||:::||| GAIGCCCACCAGGGGTGACCGGGGTGACCGGGGGTCTACGGG
                                                                                              PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThr11eGly
                                                                                                                  GlyvalThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                    ThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSerSerAlaValAsnValGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740) x US-10-156-761-1 (1-9025608)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           other or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/10156761
Publication No US20030119018A1
GENERAL INFORMATION:
APPLICANT: INEDA, HARUO
APPLICANT: ISHIKAMA, HIROSHI
APPLICANT: SHIRAA, HIROSHI
APPLICANT: SHIRAA, HIROSHI
APPLICANT: SHIRAA, TOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
ITILE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR APPLICATION NUMBER: JP 2001-272697
NUMBER: OF SEQ ID NOS: 15109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: DNA ORGANISM: Streptomyces avermitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ή
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.16e-224
2429.50
73.99%
59.03%
60.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t,
                                                                                                                                                                                                                                                                                                                                                                                               2299 GACGGC 2304
                                                                                                                                                                                                                                                                                                                                                               739 SerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 1
LENGTH: 9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-10-156-761-1
                                                                   1999
                                                                                                                                 2059
                                                                                                                                                                                                                                                               2179
                                                                                                                                                                                                                                                                                               719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                   639
                                                                                                   629
                                                                                                                                                                 619
                                                                                                                                                                                                                                 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      쉽
                                                                                                                         g
                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                 ò
                                                                 셤
                                                                                       ò
                                                                                                                                                              ò
                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1590
                                                                                                                                                1065
                                                                                                                                                                                                                  1125
                                                                                                                                                                                                                                                                                                                                    1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTGATGTACACGTCGCCGAACTTCACCACGAGCACCAGCCTCGACTACGCCGAGACCGAC 1545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TICAAGICCGGCAGGITCTACGICAGITCGGACGCGGCGACCTICACGGCGTCCGCG 1884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |||||||::|||||||::
1126 TATCCGAGCCGCTCGAACCGCTTCACCATGGATGTCTCGTCCTCGCCCTGGCTCACCTGG 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |||::::::||||||| :::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGCACCGGG-----GTGCAGTACACGACCAGGTTCGGCACCTCGTGGTCGGCGTCCGCG 1764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        638
                                                                                                                                                                                                                                                  365
                                                                                                                                                                                                                                                                                                              366 GlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAlaMet 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGly 544
                                                  SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrile 305
                                                                      GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ala---GlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      306 AspArgGlnHisProAsnThrlleMetValAlaThrGlnIleSerTrpTrpProAspThr
                                                                                                                                    TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrilePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           545 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAla
                                                                                                                                                                                326 IlellePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSer---GlyAlaProLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyVal
                                                                                                                                                1006
                                                                                                                                                                                                                                                  346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1885
                    895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        624
                                                  286
```

셤

원 장

ò

g

ద

ò

셤

ò

9 9 9 9

ò

g

δ

Д

ò

ò

g

g

ò

ò

D

g

ò

ò

g

95 G	2277568 CGCTCCGAGAAGAACCTCGCCTACGCCGCACCGACATCGGCGGCGCCTACCGCTGGGCC 2277627 46 AlaalaAsnGlyArgTrpIleProLeubeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65 ::::::	& A &	406 ThrAsnAspLeuThrLysT 2278699 ACGACAACCTCACGAACT
장 임	66 ASINGLYVALVaLSerIleAlaAlaAspProlleAsInThrAsILysValTrpAlaAlaVal 85 2277688 ACGGGTGTCGTGACCTCGCCTCCGACTCGTCGACCCGACTCGACTCGACGCGCCGTC 2277747	∂ 8	
oy B		& 4g	445 IleSerAlaLeuGlyAspl 2278819 TTCAGCGCGCTCGGTGACA
S G	125	& 8	465 ThrilePheThrSerProv :::::: 2278879 CTGATGTACACGTCGCCGA
کن [.] ط	YGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145	\$ B	485 ProSerIleIleValArgA 2278939 CCGGGCACGGTGGTGCGGG
کې لو		S &	505 ValAlaPheSerThrAspG
\ \delta \.		P &	525 ThrThrGlyGlyThrValA :::
1 & A		oy B	545 AspProGlyGlnProValV
는 상 임 상	206 SerlysThrilePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAsp 225 2278108 ACGCGGACGGTGCGTCGCCGAAGGACAACGCAACGACGACGAACGA	ò d	565 GlyValProAlaAsnAlaG ::: 2279158 GGCCTCCCGGCCGGGCGA
& £	AlaThrIrpGlnAlaValProGlyAlaProThrGlyPhelleProHisLySGly	S A	585 LeuSerAsnGlyThrPheT 2279218 TTCAAGTCCGGCAGGTTCT
l & 8	AspProValAsnHisValLeuTyrIlealaThrSerAsnThrGlyGlyProTyr	රු යු	605 AlaGlyLeuProSerS 2279278 GCCACGGGCCTGCCGAGCG
<u>ک</u> ج	SerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArglle	Cy Dp	624 GluGlyAspLeuTrpLeuA ::: :: 2279332 AAGGGGGACATCTGGCTGG
1 & E	ValProSerThrAspThrAlaAspTyrPheGlyTyrSerGlyLeuThrIle	yo B	639 ThrAsnGlyGlySerSerT :::: :::::: 2279392 ACGGACGCGGCCCT
3 8 8	31nHisProAsnThr1leMetValAlaThrGln1leSerTrpTrpProAspThr 	λ _ο qα	659 PheGlyLysSeralaProG
3 8 8	345	V da	679 GlyValThrGlyAlaTyrA
3 & 6	22/8459 TYTPTOASHANGSELEUARGTYVVALLEUASHIGSGRAGGCTIGSGACTACACCTICG 22/8518 346 TYTPTOASHANGSELEUARGTYVVALLEUASHIGSGRAAGGCTIGSGACTHATARPHE 365 378519 365 378519 365 378519 365 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 378519 37	QY Db	699 AspGlnHisGlnTyrGlyA 2279572 GATGCCCACCAGTGGGGTT
} & 4	385	yy au	719 ArgValTyrIleGlyThrA
3 & 6	386 AlalleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAla 405	QY	739 SerGly 740 2279692 GACGGC 2279697
3 4	22/8659 GAGATICGACTICGACTICGCGCGCATGATGTACGGAACGGGCGGCGACGGTCTACGGC 2278698	4 1	

NValAsnAspLeuIleSerProProSer---GlyAlaProLeu 444 ValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn 484 FyratgSerThraspGlyGlyValThrPheGlnProValAla 604 AsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaPro 738 AlaAlaSerAlaAspGlySerArgPheValTrpAlaProGly 544 SerGlyalavalGlyvalMetPheHisAlavalProGlyLys 623
::----GTGCGCTTCAAGGCGCTGCCGGCACG 2279331 ||||::||||||| |GGGGCCTCGTACCAGCGCTCTACACCAGCGCGAAGATCGGC 2279511 :::::: :::||| TTCACCAAGCTCGCCACGTCGACCAGGCCATCGGC 2279451 464 AladlySerPheAspProSerSerGlnProAsnAspArgHis 504 GlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyVal 524 ValtyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 564 GlnIleArgSerAspArgValAsnProLysThrPheTyrAla 584 AlaAlaSerSer-----GlyLeuTyrHisSer 638 TrpSerAlalleThrGlyValSerSerAlaValAsnValGly 658 GlySerSerTyrProAlaValPheValValGlyThrileGly 678 ArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 698 AsnTrpGlyGlnAlaileThrGlyAspHisAlaAsnLeuArg 718

```
:::
|||||||:::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrHisAla-----AspValThrAlaValProSerThrIlePheThrSerProVal 471
                                                                                                                                                                                                                                                                                                                                                                                     TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
                                                                                                                                                                                                                     AsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSer
                                                                                                                                                                                                                                                                                                                                                                                                               ::: ||||||| :::|||| ||| TTTGCAAAAGCCAGGAACCCTTGTTGTTGTTGTTTGTACTTTGAAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ::: ||| ||| ::: |||| ACGACCAGGAGCCACGTTCCTTTGTTGGCACCAGGTGATAACATCACTGCTTCA
                                                                              GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly
                                                                                                                                                                                                                                                                                                                                           GGAACTTGGAAAGACATCACCCCTGTCTCTGGATCTA------TACTTTGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SerTrpTrpProAspThr1le1lePheArgSerThrAspGlyGlyAlaThrTrpThrArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---ProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Arcaagcgccrccgcrcgarrgarrgagrcrcrcgagarrgacccaaccgacagcaac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGlyGlyPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       585 edecegacericas en contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra contra c
                                                      ValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly
                                                                                                                                       PhelleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp
                                                                                                                                                                             TACTTTCCTCACAAGGGGAAACTGCAGCCAGCAGAGAAGGCCTTGTATCTGACCTATTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             532 AlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GluProTrpLeuThrPheGly---ValGlnProAsnProProValProSer-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     415 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1540
                                                                                               709
                                                                                                                                       240
                                                                                                                                                                               694
                                                                                                                                                                                                                                                               829
                                                                                                                                                                                                                                                                                                                                                                                                                             940
                649
                                                         220
                                                                                                                                                                                                                         260
                                                                                                                                                                                                                                                                                                         280
                                                                                                                                                                                                                                                                                                                                               889
                                                                                                                                                                                                                                                                                                                                                                                         300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       435
                                                                                                                                                                           g
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                    ò
                                                                                    셤
                                                                                                                              ઠ
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPhelysLeuGlyGly 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIe 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITICATGGAAGAAGGTCAAGCTCGGCGGCGGCGGGGGCTTCGTCCCCGGCATCATCTTC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :::
CATCCCAAGACAAAAGGCGTAGCATATGCACGAACAGATATTGGCGGGCTGTACCGCCTC 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AACATGCCAGGACGCGGAGCCGGAGAGCGTCTGGCTGTCGATCCGGCCAACTCCAACATC 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITICCAAGGICICGITCACGCAACTGGGACGTACATCCCAGACCCGAGIGAITCC 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer 200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---AsnAsnPro 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal------GlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProlleAsnThrAsnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CGCTCGTCAGACCGCGCGCGCAACGTGGTCCTTCACCAACTTGCCCTTCAAAGTCGGGGGT
                                                                                                                                                                               and Nucleic Acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2517
325
123
264
44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-3 (1-740) x US-10-026-994-4 (1-2517)
                       Sequence 4, Application US/10026994
| Publication No. US20030113732A1
| GENERAL INFORMATION:
| APPLICANT: Dunn-Coleman, Nigel
| APPLICANT: Goedegebuur, Frits
| APPLICANT: Ward, Michael
| APPLICANT: Ward, Michael
| TITLE OF INVENTION: EGVI Endoglucanase and Ni
| TITLE OF INVENTION: EGVI Endoglucanase and Ni
| TITLE OF INVENTION: ECOI Encoding the Same
| TITLE REFERENCE: GC698
| CURRENT APPLICATION NUMBER: US/10/026,994
| CURRENT FILING DATE: 2002-04-30
| NUMBER OF SEQ ID NOS: 4
| SOFTWARE: FRASESQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.78e-145
1579.00
59.26%
42.99%
39.12%
                                                                                                                                                                                                                                                                                                                                                                TYPE: DNA
ORGANISM: Trichoderma reesei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                             2517
                                                                                                                                                                                                                                                                                                                                                                                                        US-10-026-994-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                             LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <u>ک</u> ۾
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

1239

394

414

1359

434

1419

454

1584

511

491

1644

551

571

ω

g &

Š

g 8 g ∂ g ò 셤 ò ద ð g ò 엄 ₹ d

```
||||||||:::
1160 GCACCGTGGATCAAGAACAACTTTATCGATGTGACGAGGGAGTCACCGTCCGATGGTCTC 1219
                                                                                                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                       100
                                                                                                                                                                                                                   388
                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                           448
                                                                                                                                                                                                                                                                                                     140
                                                                                                                                                                                                                                                                                                                                 508
                                                                                                                                                                                                                                                                                                                                                            160
                                                                                                                                                                                                                                                                                                                                                                                                                                               628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ileTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 359
                                :::
CATCCCAAGACAAAAGGGGTAGCATATGCACGAACAGATATTGGGGGGGCTGTACCGCCTC 220
                                                                                                 277
                                                                                                                                                          328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
                                                                                                                              80
                                                                        9
                 44
             AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp :::
                                                                                        GIGTATGCCGCAGTCGGCATGTATACGAACAGCTGGGATCCGAGTAATGGAGCCATCATT
                                                                                                                                                                                                                                                                                                                                                         LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---AsnAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ::: |||||||||||
809 TACTITCCTCACAAGGGAAACTGCAGCCAGCAGAGAAGAGGCCTIGTATCTGACCTATTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnThrGlyGlyBroTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               929 GGAACTTGGAAAGACATCACCCCTGTCTCTGGATCAGATCTA-----TACTTTGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TCTTGGTGGCCAGATGCTCAGCTGTTTCGGTCGACCGACTCTGGGGACAACATGGAGCCCG
                                                                                                                              AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsapProIleAsnThrAsnLys
                                                                                                                                                                                     ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu
                                                                                                                                                                                                                                                            AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAspAsnIle
                                                                                                                                                                                                                                                                                                                     TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                 TITICCAAGGICICGICGITCACGGCAACIGGGACGIACAICCCAGACCCGAGIGAIICC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerTrpTrpProAspThr1le1lePheArgSerThrAspGlyGlyAlaThrTrpThrArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Archadacticacadactarccaactgaacctartactacagcarctcaactccaaa
                                                                AspAlaAlaAsnGlyArgTrplleProLeuLeuAspTrpVal-------GlyTrp
                                                                                                                                                                                                                                               ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PhelleProHisLysGlyValPheAspProValAsnHisValLeuTyrlleAlaThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             360 GluProTrpLeuThrPheGly---ValGlnProAsnProProValProSer----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       340
                                                                                                 221
                                                                                                                                                                                                                                                                         389
                                                                                                                                                                                                                                                                                                                                                                                                                                               569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                280
                                                                        45
                                                                                                                              61
                                                                                                                                                                                       81
                                                                                                                                                                                                                   329
                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                449
                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                      509
                                                                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320
                25
                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                         8 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                       역 상 역
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 %
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
8
                 8 6
                                                                                        g
                                                                                                                                                  g
                                                                                                                                                                              ₹
                                                                                                                                                                                                                                                                                                                       8 8 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                     ò
                                                                                                                           \delta
                                                         1815
                                                                                                                                                                                                                                                                                                                                                                                        2092 GGAGACAGCGGCCTCCTGGACGGACATCCAGGGCTCCCAGGGCTTCGGCTCCATCGAC 2151
                                                                                                                                                                                                                                                                              2091
                                                                                                                                                                                                                                                                                                                                                                                                                              CAGCGCTCGCAGTTCCAGGCAGCTTTGCCTCCGTCTCGAGCCTGCCCGCGGGCGCCGTC 1755
                                                                                     611
                                                                                                                                         612 AlaValGlyValMet-----PheHisAlaValProGlyLysGluGlyAspLeuTrpLeu 629
                                                                                                                                                                                                                                                                                                                   686
                                                                                                                                                                                                                                                                                                                                                                         706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         630 AlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---
                                                                                                                                                                                                                   ---AACCTGTATGCCTTCGGCACC-----GGCCCGTCAGGGGCTCGCCTCTACGCCAGT
                                             |||| |||||||::
ATCGCCTCGGACAAGAACCAACAGCGTCTTCTACGCCGGCTCCGGATCGACCTTTTAC
                                                                                                    649 ThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr
                                                                                                                                                                                                                                                                                                                ProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla-----TyrArgSer
                                                                                                                                                                                                                                                                                                                                                                        AspaspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrp
                                                                                  592 ArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGly
                            IleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ArgGlyIleValTyr-----GlyAspIleGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Dunn-Coleman, Nigel
APPLICANT: Goedegebuur, Frits
APPLICANT: Ward, Michael
APPLICANT: Yac, Jian
TITLE OF INVENTION: EGVI Endoglucanase and Nucleic Acids
TITLE OF INVENTION: Encoding the Same
TITLE OF INVENTION: Encoding the Same
FILE REFERENCE: GC698
CURRENT APPLICATION NUMBER: US/10/026,994
CURRENT FILING DATE: 2002-04-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2710
325
123
264
44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                x US-10-026-994-1 (1-2710)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application US/10026994 Publication No. US20030113732A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1.97e-145
1579.00
59.26%
42.99%
39.12%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: DNA
ORGANISM: Trichoderma reesei
US-10-026-994-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-3 (1-740)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 2710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-10-026-994-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
```

셤

ઠે

ò D

ద ò g g 장 g 8 ద ò ద ò q 8 셤 ò g ò 셤 ò

ð

```
GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProserGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlalleLeuArgSerSerAspGln 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 TyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46 AlaAlaAsnGlyArgTrplleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   421 dececcácineda consecutor contra contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de contra de 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
2217
308
104
257
86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-3 (1-740) x US-10-156-761-2561 (1-2217)
                                                                                                                                                                                                                    APPLICANT: ONURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: HATTORI MASAHIRA
ITILE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFRERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT PILING DATE: 2002-05-29
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 2561
                                                                                                                                  ; Sequence 2561, Application US/10156761; Publication No. US20030119018A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: DNA ORGANISM: Streptomyces avermitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.17e-128
1405.00
54.57%
40.79%
34.81%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .. (2217)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; LOCATION: (1).
US-10-156-761-2561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                               2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1855
                                                                                                                                                                                                                                                                                                                                        2131
                                                                                                                                              1685 TATTCGGCCGACGCGACACGATCCTCTGGTCGACCGCCTCGTCCGGC------GTG 1735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1220 ATCAAGCGCCTCGGCTGGATGATTGAGTCTCTCGAGATTGACCCAACCGACAGCAACCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1625 GGCGCGACGTGGAGCATCGACTACGCGGCCGACACGTCCATGAACGGCGGCACGGTGGCC 1684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGCACCAAGGTCGCCGGCAGCGGCAGCACCGCGGGCAAGTCTACGTGGGCACCAACGGC 2251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          532 AlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValVal 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   989
                                                                                                                                                                                                                            GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAsp 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            552 TyralavalGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGln 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            592 ArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGly 611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    612 AlaValGlyValMet -----PheHisAlaValProGlyLysGluGlyAspLeuTrpLeu 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr 668
                                                                                                           395 MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrHisAla------AspValThrAlaValProSerThrIlePheThrSerProVal 471
  ---ProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProhlavalPhevalvalGlyThr11eGlyGlyvalThrGlyAla-----TyrArgSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1796 Arcecerceakchagaagaccaacageerririacececececececececeririiac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---AACCTGTATGCCTTCGGCACC-----GGCCCGTCAGGGGGCTCGCCTCTACGCCAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              687 AspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlyGlnAla1leThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ArgGly11eValTyr----GlyAsp11eGlyGlyAlaProSerGly 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      415 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                699
                                                                                                                                                                                                                                                                                                                                                                                                                                                            455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2081
     376
```

g ò g ò g ò g ∂ g 8 g ò

480

363

85

420 125

243

303

65

183

45

512 GlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThr 1525 GGGCGACCTGGGCCCCTTCGAGGCCCAGCCCGACATCGCCAAGGACGCACCG 530 ValAlaAlaSerAlaAspGlySerArgPheValTrpAla	Cy 562 AlaSerGinGlyValProAlaAsnAlaGinIleArgSerAspArgValAsnProLysThr 581. 1690 GAGGTCTCCTCCCGAAGGGCGCCACCGGTCGCCGACCGGCCGATCCGACCGGC 1749 Cy 582 PheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyCalThr 599 [620 ValProGlyLysGluGlyAspLeuTrpLeualaAlaSerSerGlyLeuTyrHisser	Oy 679 GlyvalThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 2046 Oy 679 GlyvalThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 698 1:::	Qy 719 ArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAsp 733	APPLICANT: OWURA, SATOSHI APPLICANT: INEDA, HARUO APPLICANT: ISHIKAWA, JUNO APPLICANT: SHIRAWA, HROSHI APPLICANT: SHIBA, TADAYOSHI APPLICANT: SHEBA, TADAYOSHI APPLICANT: SHEBA, TADAYOSHI APPLICANT: SAKAKI, YOSHIYUKI APPLICANT: APPLICANTON: NOVEL POLYNUCLEOTIDES ITILE OF INVENTION: NOVEL POLYNUCLEOTIDES CURRENT APPLICATION NUMBER: US/10/156,761 CURRENT FILING DATE: 2002-05-29 FRIOR PELLING DATE: 2001-05-30 PRIOR PELLING DATE: 2001-05-30 PRIOR APPLICATION NUMBER: JP 2001-272697 FRIOR PILING DATE: 2001-06-30 PRIOR APPLICATION NUMBER: JP 2001-272697 FRIOR PILING DATE: 2001-06-30 SRO ID NO 1 SEQ ID NO 1 LENGTH: 9025608 TYPE: DNA
166 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185	PheTrpSerArgAspGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe :::	278 ThrSerGaCGGGCCGGGCGGCGGCGGCGGCGGCGGCGGCGCGCGC	1012 GTGTCTCCACCACACCACCAGGGCCGACGCGACACGGTCTTCCGGTCCGGACGGC 1071 334 GlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyr 353 1072 GGCCGTACCTGGACGTCCCTCAAGGAC	ProSerProLysLeuGlyTrpWetAspGluAlaMetAlaIleAspProPheAsnSerAspAAGCCGAAGTTCGGCTGGTGGTCCAGGCGCTCGGGTCGACCCGTACGAG ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLySTrpAsp :::::	414 SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsn 433 1264

```
3160584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3160421 ATCGCCACCAGTGCCGACGGCGCACACTGCTGGTCCTTCGTGCACTGGGACGGC-- 3160365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3160140
                  :::|||:::|||
GTCGTCTCCACCACCAACAACAGGCCGACGACGACGGTCTTCCGGTCCACGAGGGC 3160935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3160482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3160257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          638
                                                                 313
                                                                                                                                333
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsn 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   511
                                                                                                                                                                                                                                                                                                                                           3160583 TTGGGGTGACGGACTCGCGCAGGCGGCCGGCCAGCCGGCGTATGTCGTCCGCACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3160316 GAGGTCTCCTTCCCGAAGGGCGCCAACCCGGTCGCCGACCGGCCGATCCGACGCGC
                                                                                                                                                                                          GlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GlyLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThrThrGlyGlyThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -- ProGlyAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------ACGACGTACGCGGCCCACCGCTCGACGGACAACGGCGCGAGCTGGTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3160256 Tricracedecraceacrregacaaresaacecraraceccasecacreacaereses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3160196 Trcacescocordedescodacrece---recesceacascocastrcaascreses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        620 ValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer---SerGlyLeuTyrHisSer
                                                                                                                                                                                                                                                                                                                                                                                          ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProGlyGlnProValValTyrAla -----ValGlyPheGlyAsnSerTrpAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThr
                                                               AlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIle
                                                                                                                                                                                                                                                                                                                           ProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3160523 GGCTGGGGCGACCACGGCAACGGC--------GCGTACTCCCACGACGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAla
                                                                                                                           MetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGly
                                                                                                                                                                                                                                                          ValleuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       434 AspleulleSerProProSerGly -- AlaProLeulleSerAlaLeuGlyAspLeuGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyPheThrHisAlaAspValThrAlaValProSerThr---IlePheThrSerProVal
                                                                                                                                                                                                                                                                                                                                                                                                           3160799 CACGTCGTGTACGGGACCGGCGCACCTCTACGGCACCCGCGACCTCAAGCGCTCGAC
                                                                                                                                                                                                                                                                              ValalaalaSerAlaAspGlySerArgPheValTrpAla--
                                                                                                                                                                                                                  3160934 GGCCGTACCTGGACGTCCCTCAAGGAC------
278 ThrSerGlyThrTrpThTArglleSerProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                         414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              009
                                                                                                                                                                                                                                                                                                                            374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453
                                                                                                                                                                                               334
                                                                                                                                                                                                                                                               354
                                                                                                                                                                                                                                                                                           3160901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
G
                                                                                                                                                      g
                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                 8686
                                                                                                                                                                                                                                                                                                                                                                                8 6 8 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                장염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  $ B $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                         g
                                                                 유
                                                                                                                    8
                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3161882 TACCGCTGGCGAACGCCGTCATCGGGGGCACCGGCTTCGTCACCGGCGTGCTCTTCCAC 3161823
                                                                                                                                                                                                                                                                                                                                                                                                                             3161762 GACCGCGGCCCGCTGGACCCCCGCTCATCGACCTCGGGACGACGACGACGTC 3161703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3161642 GGCACCTACGCCCAGTCGTGGGCGGCAAC---GGCGGGGTCCTGCGCTCCGAGGACCGC 3161586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGCGCCACCTGGACCCGCACCGGACCTGACCGTGAAGCTCGGCGGCAACGAGGACGGCCGC 3161526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GCGAAGGCGAACTCC 3161385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGCGCACCGTCTACGCCGGCTGGGGTGACGGCGACGCACCTCGGGCACGGCGAACCTG 3161292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCCGCCAAGGTCCCGCTCCGCGCGTACGACACGCACACCCGCGAGCTGTACGTGACG 3161175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SerLysThrIlePheValGlyValAlaAspProAsnAsnPro----------Val 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------IleproHisLysGlyValPheAspProValAsnHisValLeuTyrIleAla 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82
                                                                                                                                                                                                                                                                                                                                           45
                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                              AspIleGinGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAla
                                                                                                                                                                                                                                                                            rThrTrpSerAsnValAlaileGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                          GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                           AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AsnGlyvalValSerIleAlaAlaAspProIleAsnThrAsnLy8ValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerVal
                                                                                                                              9025608
308
104
257
21
                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                              x US-10-156-761-1 (1-9025608)
                                                                 or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TCCGGGCAGGGAGTCGTGTTC-----
                                                                                                                                                                                                                Gaps:
                                                                  other
   avermitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3161408 GCCTTCCCG------
                                                                    ή
                                                                                                                              1.24e-123
1405.00
54.57%
40.79%
                                                                  ΰ
                                                                                                                                                                                               34.81%
                                                                 ú
 ORGANISM: Streptomyces FEATURE:
                              NAME/KEY: misc feature;
LOCATION: (4187715);
OTHER INFORMATION: a, t
US-10-156-761-1
                                                                                                                                                                                                                                              US-09-917-376-3 (1-740)
                                                                                                                                                                                Best Local Similarity:
                                                                                                                                                                Percent Similarity:
                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     166
                                                                                                                                                                                                                                                                                                                                                                           3161822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3161702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3161384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3161351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3161585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3161465
                                                                                                                                                                                              Query Match:
                                                                                                                                . No. :
```

g

ò 셤 ò

셤 à 요 . 장

à g ò 셤 8 8

ò 셤 ò 유 ò a

8 원 à g à

g

ð

:::	qu	:::::: ::: 1457 CIGTATCIGGCGGCGGCACCIAIAIGCAIGAACGCGCCGGCACTGCCGCAGIGIIG 1513
Oy 639 ThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSerSerAlaValAsnValGly 658	yy Q	101 ArgSerSerAspGlnGlyAlaThrTrpGln1leThrProLeuProPhelySLeuGlyGly 120
Qy 659 PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrileGly 678	Qy	121 ASIMETProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsIASpAsII1e 140
Oy 679 GlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 698 :::	oy Oy	
Qy 699 AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 718 Db 3159899 GACGCCCACCAGTGGGGGGGGGGGGGGGGGGGGGGGGGG	ò a	161 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
Qy 719 ArgValTyrileGlyThrAsnGlyArgGlyIleValTyrGlyAsp 733	S da	181 ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer 200 1751 GTTGGGCGCGAGCAGGCGGTGGGGATCGCCTTTGTCGTGTTCGACGCAGTGGC 1807
RESULT 9 US-09-927-827-21 Sequence 21, Application US/09927827	S da	201 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal 220 :::
	S da	221 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240 :::
; IIILE PEFERENCE: 38-10(15824)B ; CURRENT FILING DATE: 2001-08-10 ; CURRENT FILING DATE: 2001-08-10	S qa	241 IleProhisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
FALOK FULLON NOMBEK: US 60/2/9,493 PRIOR FILING DATE: 2001-03-28 NUMBER OF SEQ ID NOS: 69 SEQ ID NO 21	Qy QD	261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 280
TYPE: DNA ORGANISM: FEATURE:	ζς Op	281 ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 300
; NAMBL/KET: CDS ; LOCATION: (1001)(2668) US-09-927-827-21	oy Og	301 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 320 :::::
Alignment Scores: 7.17e-102 Length: 3668 Score: 1139.00 Matches: 267 Bercent Similarity: 50.85\$ Conservative: 120	yo da	TrpTrpProAspThr11e11ePheArgSerThrAspGlyGlyAlaThrTrpThrArg11e
uery Match: 28.22% Indels: B: 10 Gaps:	ò 4	341 TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 360
US-09-917-376-3 (1-740) x US-09-927-827-21 (1-3668) Oy 1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaileGlyGlyGlyGlyPheValAsp 20	G &	ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp
	名 &	2255 CCGTGGACCGCACATGCCACGCCGCAC
Oy 21 GlylleValPheAsnGluGlyAlaProGlylleLeuTyrValArgThrAspIleGlyGly 40	g q	ATGGGGGCGCTGGCGATCGTTCGACGGCAACCATGCGCTGTTCGTGACCGGC
31yTrp 	& a	401 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle 420 11.:::: ::: ::: 2342 TACGGCATCTGGGCCTCGCGCAATCTG
61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80	S G	421 AlaPro
Db 1397 GACGACTGGAACCTGATGGGCATCGACGCGTTCGCCGTCGACCCGCCGATGCCGATGCG 1456 Qy 81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100	۸ŏ	433 AsnAspLeulleSerProProSerGlyAlaProLeulleSerAlaLeuGlyAspLeuGly 452

원

ò

g à g ò g 8 셤 δ

ò

```
Indels:
                                                                                                                                                                                                                                                                                                                                5.94e-100
                                                                                                                                                                                                                                                                                                                                          1117.00
50.37%
34.95%
27.68%
                                                                                                                                                                                                        NAME/KEY: mat_peptide
LOCATION: (1)...(2367)
FEATURE:
                                                                                                                                                                                                                                                                 .. (2367)
                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                      NAME/KEY: CDS
                                                                                                                                                                                                                                                                    LOCATION: (1)
US-10-395-241-13
                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              165
                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ⋩
                                                                                                        GCGCACGTCCGACTTCGGCGCGCGAGGGTGGTGAACGCCCTGCCGAACCCCGGGTG 2848
                                                                                                                                                                                                                                                                                                                                                                                                         3284 TCCGCGCATTGCCGGGCGTGTGTACTTCGCCACCGGCGCCGCGGCATTTTCTACGGCGA 3343
CTGGACCTGCTCAGCCCGATGGCTGGCGCGCATCTGCTCAGCGCGCTCGGCGATATCGAC 2500
                                                                                                                                                                  :::
ACCGTGCGCCCGCCGCAACGAAATCCGCGCGC---TGTATTACGCGTGATGGCGG 2674
                                                                                                                                                                                                                                     CAAGCAATGGACTGCGTTTGCGAGTGAGCCGCCAGCCGGCGGCGCCGGCGAGCATTGC 2734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           713
                                                                                                                                                                                                                                                                                                                             551 lTyralaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGl 571
                                                                                                                                                                                                                                                                                                                                                                                        nlleArgSerAspArgValAsnProLysThrPheTyrAla-----LeuSerAsnGlyTh 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAs 733
                                                                                       492
                                                                                                                                                                                                        ylysAsnTrpPheGln---GlySerGluProGlyGlyValThrThrGlyGlyThrValAl 531
                                                                                                                                                                                                                                                                    aAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValVa 551
                                                                                                                                                                                                                                                                                                                                                                                                                                                 rPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeu---- 607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ralaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVa 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSerSe 653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    613 lGlyValMetPheHisAlaValProGlyLySGluGlyAspLeuTrpLeuAlaAlaSerSe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                673 lvalGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrTr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAs
                             GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPhe
                                                                                       ThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGly
                                                                                                                                                 SerPhe-AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3023 AGGCGTGGTGTATCTGGCCAGCCCCGGCAAG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-10-395-241-13
Sequence 13, Application US/1031;
Publication No. US20040038367A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: YAOI, Katsuro
APPLICANT: MITSUISHI, Yasushi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           T 3344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              p 733
                                                                                                                    2558
                                                                                                                                                                             2618
                                                                                                                                                                                                        512
                                                                                                                                                                                                                                       2675
                                                                                                                                                                                                                                                                    531
                                                                                                                                                                                                                                                                                                 2735
                                                                                                                                                                                                                                                                                                                                                                                                                                                 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3344
 2441
                                                                                     473
                                                                                                                                                 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          653
                             453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
```

ద

엄

8 셤 ò

ò

ద

8

셤

ò g ò 셤 ò g 8 g 8 q 8

```
TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLECTIDE FITLE REFERNCE: 073756 CURRENT APPLICATION ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME CURRENT APPLICATION NUMBER: US/10/395,241
PRIOR APPLICATION NUMBER: US/203-03-25
PRIOR APPLICATION NUMBER: US/202-83433
PRIOR APPLICATION NUMBER: US/202-83433
PRIOR PILING DATE: 2002-03-25
NUMBER OF SEQ ID NOS: 19
SEQ ID NO 13
LENGTH: 2367
TYPE: DNA
ORGANISM: Geotrichum sp. M128
FEATHER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            105 GlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGly 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGln 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            205 AlaSerLysThrilePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||::: |||||||
| ACGICCAICCCGGACGCTIC------ACCAACGGTAICGGATACACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 ATGGGCACCGAGTCGATCGCGCTGGACCCCAACAACCCCGACAGGCTGTACCTCGCGCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TyrThrTrpSerAsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 TACGAGTTCAAGAATGTCGCGATCGGCGGCGGCGGCTACATTACCGGGATTGTCGCGCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     70 ccaaagaccaagaccrocrororataccococococarroscocococoraccoroscocoros com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlyMetTyrThr---AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             250 edececrarerceccaccaccacreses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          125 ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      145 AlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         118 ACGCGTACA---GAGGGTATCTGGAAGAGTTCGGACCGCGCCCAAGACCTGGACAAACGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                185 SerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2367
281
124
279
120
31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-917-376-3 (1-740) x US-10-395-241-13 (1-2367)
```

46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65 133 GCAGGCACGTCCAAGTGGATCCCCTCAACGACTTTATCGAGGCGCAGGACATGAACATT 192 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85 193 ATGGCACCGATCGGTTGGACCCCAACACACCGGCAGGACGTGCTCTGGGAG 252 86 GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAsp 104		361 CGCAACAATGGCGAGCGCTCGCTGTCAACTCGAACGAGGGGGTTGGAT 420 145 AlaProSerGlyLygGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet 164 121 ACGCGTACAGAGGGTATCTGGAAGATTCGGACCGCGCCCAAGACTTGGAAAACTTCGAACTGGAAAACGTC 477 165 ThrAsnPheProAegyValGlyThrTyr1leAlaAsnProThrAspThrThrGlYTyGln 184	ACGTCCATCCCGGACGCGTTCACCAACGGTATCGGATACACG SerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln	200 ICG	AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle	247 PheAspProValAsnHisValleuTyrIleAlaThrSerAsnThrGlyGlyProTyrAsp 266	GlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArglle 2 :::	832 CCCGGGGGGGACTGGCGGCGGTACAACAACCAGAGGTTCCCTGGGGGGGA 891 300 TyrSerGlyLeuThrIeAspargGlnHisProAsnThrIeMetValalaThrGlnIle 319 11 11 11 11 11 11 1	949 GACCGCGACCCCGCCCTCGACACACATCTACCTCTCAACCGATGCCGGCGCGACC 1008 337 TrpThrArg11eTrpAspTrpThrSerTyrPro 347	348 AsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVal 367
8 8 8 8 8	OY DP OY	8 8 8	8 8 a	8 & 8	\$ 6 \$ 5 *	a va	ó a ó	6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7	8 & B	장 연 .
Oy 619 AlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHis 637 Db 2106 GTGAACCCTGGGTGGCCGGCGGCTCTTCCCGAGGGGGTCTTTCCAC 2165 Oy 638 SerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsn 656 Db 2166 TCGACCGATTTGGCGCTCGTTCACGAGGGCCCCAACGCGACGAGGCCCCCACGAGGCCCCACGGACGCCCCCACGAGGCCCCCAACGCCCAACGCCACCTCAGGCCCCCAACGCCCAACGCCACCTCAGGCCCCCAACGCCCACCCA	Qy 673 ValvalGlyrhrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGly 690 1	ccredacgcgcgrcaardacag	Db 2463 TACGCCGACCTT 2474 RESULT 12 US-10-395-241-17	-	1 TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHAKIDE-DEGRADING ENZYME, POLYNUCLEOTIDE 1 TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME; FILE REFERENCE: 073756; CURRENT APPLICATION NUMBER: US/10/395,241; CURRENT FILING DATE: 2003-03-25; PRIOR APPLICATION NUMBER: JP 2002-83433; PRIOR PILING DATE: 2002-03-25; NUMBER: JP 2002-83433	SOFTWARE: Patentin version 3.2 SEQ ID NO 17 LENGTH: 2481 TYPE: DNA	FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FORTURE: FOR	Alignment Scores: 1.25e-99 Length: 2481 Score: 1114.00 Matches: 281 Percent Similarity: 50.25\$ Conservative: 123 Best Local Similarity: 34.95\$ Mismatches: 280 Query Match: 13 Gaps: 31	US-09-917-376-3 (1-740) x US-10-395-241-17 (1-2481) Qy	26 GluGlyalaProGlylleLeuTyrValargThrAspIleGlyGlyMetTyrArgTrpAsp 73 CCAAAGACCAAGGACGGGGGCACGGACGGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC

Db 2161 AGCACCTGGACGCGCGTCAATGACCAGAACAACTACTCGGGCCCCACCATGATC 2217	Oy 711 ThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArleVal 730 Db 2218 GAGGCCGACCCAAGGTCTACGGGCGCGTGTATCTAGGCACGAACGCCGCGGTATCGTG 2277 Oy 731 TytGlyAspIle 734	Db 2278 TACGCCGACCTT 2289 RESULT 13 US-10-420-191-1 ; Sequence 1, Application US/10420191 ; Publication No. US20040067569A1	; GENERAL INFORMATION: ; APPLICANT: No. USZUO40067569Alozymes Biotech, Inc. ; APPLICANT: Rey, Michael W. ; APPLICANT: Zaretsky, Elizabeth J.	; APPLICANT: Haas, Jeffrey A. ; TITLE OF INVENTION: Polypeptides Having Xyloglucanase Activity And Nucleic Acids ; TITLE OF INVENTION: Borcoding Same ; FILE REFERENCE: 10210.200-US	; CURRENT APPLICATION NUMBER: US/10/420,191 ; CURRENT FILING DATE: 2003-04-18 ; PRIOR APPLICATION NUMBER: US 60/373,987 ; PRIOR FILING DATE: 2002-04-19	; NUMBER OF SEQ ID NOS: 6; SOFTWARE: PatentIn Version 3.2; SOFTWARE: PatentIn Version 3.2; LENGTH: 5698	; TYPE: DNA ; ORGANISM: Tricoderma reesei US-10-420-191-1	6.24e-76 Length: 878.50 Matches: 29.92% Conservati	larity: 22.06* Mismacches: 21.77\$ Indels: 13 Gaps:	-09-917-376-3 (1-740) x US-10-420-191-1 (1-5698) 7 ThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn :::	≨ ;	302 AGTACCTTCTTGCAGTTCGAGCCGCCGCCGCCGAGCAGGGGCCGTAGTAGAAGGTT 30GlylleLeuTyrValargThrAspileGlyGlyMetTyrArgTrp :::	362 CCCAAGACAAAAGGCGTAGCATATGCACGAACAGATATTGGCGGGCTGTACCGCCTCAAA 45Aspala 1::	422 GGGTTCTGTTTTCCGCATCGTATACGTGCTTGTCTATAACGGCCCGACATGGCGGGGGTTC 47 AlaasnGlyargTrplleProLeuLeuAspTrpValGlyTrp	э <u>н</u>	Db 542 CGGCTGCTGAGTACCTGGCGGCAGTGCCTACCGTACGACTATTACGGCCGACCGTGTTC 601 Qy 66 AsnGlyValValSerIleAlaAlaAspProlleAsnThrAsnLySValTrpAlaAlaVal 85	Db 602 TGGGGCATCGACGCTGTTGGTCCGCAGGACGATCAAAAGGTGTATGCCGCAGTG 661
368 GlnProAsnProProValProSerProIysLeuGly 379	O TrymetAspGluAlaMetAla11eAspProPheAsnSerAspArgMetLeuTyrGlyThr 399 O TrymetAspGluAlaMetAla11eAspProPheAsnSerAspArgMetLeuTyrGlyThr 399 O TrymetAspGluAlaMetAla11eAspProPheAsnSerAspArgMetLeuTyrGlyThr 399 O TrymetAspGargacTraTcGArgCCCTrCArgCCCGAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATG		ValAsnAspLeuileSerProProSerGlyAlaProLeuileSerAlaLeuGlyAspLeu :::	452 GlyGlyPheThrHisAlaAspValThrAlaValPrOSerThrIlePheThrSerProVal 471	472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491 	492 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 511 1519 GGATCCTCGGGACACGAGTACGACGCGTGCGCGCGCGCGTGCGT	512 GlyLysAsnfrpPheGlnGlySerGluProGlyGlyValThrThr 526 	GlyGlyThrValAlaAlaSerAlaAepGlySerArgPheValTrpAlaProGly 5	545 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 564 ::: ::: ::: :::	565 GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsn 578	579 ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVal 598 	599 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis 618 	619 AlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHis 637 	638 SerThrAsnGlyGlySerSerTrpSerAlalleThrGlyValSerSerAlaValAsn 656	657 ValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPhe 672	673 ValvaldiyThrIleGiyGiyValThrGiyAlaTyrArgSerAspAspCysGly 690 	ThrThrTrpValLeulleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlalle

	156 AspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAla 175	3625 CCGGGCACCCCGCTGGCC 3642	176 Asn	3643 AACGGCACGCAGGCCCACGGCCACCGACCGGCAATACCGGCCGCAGGCC 3702		3703 GCCACCACGGTGGACGCGGTGGCGCCGCCGGTGATCGATC	202 LeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsn 218	3763 ACCATCAGCGGCACCGGGGGCCGGGGCCAAGGTGATCCTCACCGACGGCAACGGCAAC 3822	219 ProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro 237	3823 CCGATCGGCGAAACCACCGCCGACGCGCAACTGGAGCTTCACGCCGGCACGCCG 3882	238ThrGly 239	3883 CTGGCCAACGGCACGGTGACGCCGTGGCCCAGGACCCTGCGGGCATACCGGCCCG 3942	240PhelleProHisLysGlyValPheAspProValAsn 251	3943 CAGGGCAGCACTACCGTGGACGCGGTGGCGAACACGCCTGTGGTCATCCGAGCAAC 4002	252 HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySer 268	4003 GGCAACCTGCTCAACGGTACGGCCGGGCAGCACCGTGACCTTGACCGAACGCAAC 4062	269 SerGlyAspValTrpLySPheSerValThrSerGlyThrTrpThrArglle 285	4063 GGCAACCCGATCGGCCAGACCGCCGGTGGCACCGGCAACTGGGAGCTTCACGCCCGGC 4122	286 SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSer 301	4123 TCGCAACTGCCCAACGGCACCGTGGTCAACGTGACCGCGGCAGCGACGCCGCCGGCAATACC 4182	GlyLeuThr1leAspArgGlnHisProAsnThr1leMetVal	AGCCTTCCCGCTACCACGACGGTGGATTCCTCGCTGCCGTCGATCCCGCAGGTGGATCCG	316AlaThrGlnIleSerTrpTrpProAspThrileIlePheArgSer 330 	のできないには、1 にもられないになってもないできない。	ACCGATGGCAACGGCAACCCGATTGGCCAGGTCACCGCCGACGGCAGCGGCAACTGGTCC	343 TrpThrSerTyrProAshArgSerLeuArgTyrVal 354		355 LeuAspileSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProPro 372	4417 AGCAATGTGGACAGTGGCGGGGGGTGATCACTGTGGATGGCGTGGCCCCCGGCGCGCCG 4476	373 ValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSer 392	4477 GTGATCGATCCGAGCAACGGC 4497	AsnAspLeu 409	45	410 ThrLysTrpAspSerGlyGlyGlnlleHislleAlaProMetValLysGlyLeuGluGlu 429
qa 	δ	qa	λö	q ₀	λõ	අු	λō	අ	ζÓ	අු	δ	d d	δ	qa	λō	q	λo	අධ	λŏ	ą _C	ài	g G	ò 4	3 8	q _Q	ò	qa	ò	qq	ò	Ωp	ò	셤	ò
; APPLICANT: Xu, H. ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms . Fits Depressive: Elimpa Alla	CURRENT PILING DATE: 2001-02-20	明ら	APPLICATION N FILING DATE:		APPLICATION NUMBER: FILING DATE: 2000-09		APPLICATION NUMBER: FILING DATE: 2000-10	; PRIOR APPLICATION NUMBER: 60/253,625 ; PRIOR FILING DATE: 2000-11-27	APPLICATION N FILING DATE:	APPLICATION N	APPLICATION N	ning Prior Application	RE: Patentin ver	; Car To Journ 1 407 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2 French 2	; ORGANISM: Pseudomonas aeruginosa ; ORGANISM: Pseudomonas aeruginosa ; ORGANISM: 2015	4 (Alignment Scores: Pred. No.: 1.27e-12 Length: 7407 Score: 238 00 Matches: 206	t Similarity: 32.59\$ Conservative:		-282-122A-30151 (1-7407	lAlaIleGlyGlyGlyGlyPheValAspGly	Db 3250 GTCAACCTGAGCAATGGCAGCAGCTCAGCGGCACGCCGAGCCGGGG 3297	Oy 31 IleLeuTyrValargThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArg 50	Db 3298 AGCACGGTGATTCTCACCGACGGCAACGGCAAT 3330	51 TrplleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr	3331 CCGATCGCCGAGGTCACCGCCGACGGCAGCAACTGGACCTGCACCGCCG	ASNGIYVALVALSETILEALAAJAABPKTOLIEASNINIASNILYSVALITYD 82 ATCGCCAACGGTAACTGTAACTGAACGGAACGAACGAAACGAAAAAAAA	5571 AICOCCRACOGIACIGGICHACOLOGGICCCCAGGACCICCCGIHACAGCAGCCC SA BIBBIANA GABANARA SA SA SA SA SA SA SA SA SA SA SA SA SA	ATAMIA VALGA COGLOGATICOS ::: CCGGCGA CGGTGA CCGTCGATICCA	97 GlvalatleToranarasiscons and ambancarasiscons an	3511 GGCGTCGTCATCACCGCCCCGAGCCCCGTGCCACACACAC	116 PheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspPro	Db 3562GATGCCGGCGAACCCGATAGGGCAGGTCACCGCCGAC 3600	Oy 136 AsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThr 155

Oy 722 IleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly 740	; Publication No. US20030166030A1 ; GENERAL INFORMATION: ; APPLICANT: O'Toole, George A. ; APPLICANT: Mah, Thien-Fah ; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF	; TITLE OF INVENTION: BIOFILM-SPECIFIC ANTIBIOTIC RESISTANCE ; FILE REFERENCE: 14537-002001 ; CURRENT APPLICATION NUMBER: US/10/246,330 ; CURRENT FILING DATE: 2002-09-18 ; PRIOR APPLICATION NUMBER: US 60/323,241	; PRIOR FILING DATE: 2001-09-18 ; NUMBER OF SEQ ID NOS: 32 ; SOFTWARE: FastSEQ for Windows Version 4.0 ; SEQ ID NO 3	; LENGTH: 7407 ; TYPE: DNA ; ORGANISM: Pseudomonas aeruginosa ; FEATURE:	; NAME/KEY: CDS ; LOCATION: (1)(7404) US-10-246-330-3	nt Scores: 1.27e-12 Length: 238.00 Matches:	Percent Similarity: 32.59% Conservative: 80 Best Local Similarity: 22.80% Mismatches: 314 Query Match: 5.90% Indels: 292 DB: 6aps: 50	-09-917-376-3 (1-740) x US-10-246-330-3 (1-7407)	OY 11 ValAlaleGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly 30 12	Qy 31 IleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArg 50	3298 AGCACGGTGATTCTCACCGAC	Qy 51 TrpileProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65		Db 3391 ATCGCCAACGGTACTGTGGTCAACGTGGTGGCCCAGGACGCCTCCGGTAACAGCAGCCCG 3450 Qy 83 AlaAlaValGlyMetTyrThrAsnSerTrpAspproAsnAsp 96	Db 3451 CCGGCGACGGTGACCGTCGATTCCAGCGCGCCGCCGCGCGCG	Qy 97 GlyalaileLeuArgSerSerAspGlnGlyAlaThrTrgGlnIleThrProLeuPro 115	116 PheLysLeuglygllydsmMetProglyArgGlyMetGlyGluArgLeuAlaValAspPro 13	Db 3562GATGCCGGCAGCAACCCGATAGGGCAGGTCACCGCCGAC 3600	Oy 136 AsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLeuTrpArgSerThr 155	Db 3601	
	Db 4645 GCCGTGGCCCAGĠĊCGGCCĠĠĊAATÁĊĠAGCGCTCCGGCCAGCGTCGATGCC 4704 Qy 460 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer 477	Oy 478 ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAapproSer 497	Qy 498 SerGinProAsnAspArgHisValAlaPheSerThrAspGlyGlyLyBAsnTrpPhe 516	Qy 516516 Db 4852 ACGCCCGGCACGCCGACGCCCAACGGCTCGGTGATCAATGCGCTGGCCCAGGACGCCGCC 4911	Qy 517GlnGlySerGluBroGlyGlyValThr	Qy 526ThrGlyGlyThrValalaAlaSerAlaAspGlySerArgPheValTrp 541	Qy 542 AlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer 559 Db 5032 CTCACCGACGGCAACGGCAACGCGATCGGCCAGTCACCGCCGATGGCGGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGAACGCAACGCAACGAACGCAACGCAACGAACGAACGCAACGAACGCAACACAACGAACGCAACACAACA	560 TrpalaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro	Db 5089 TGGAGCTTCACGCCCGGCACGCCG	Db 5113CTGTCCAATGGCACGGTGGTCAATGCGGTGGCCCAGGACGCTGCC 5157	593SerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuPro	Db 5158 GGCAACACCAGGGCCGGTCAGCACCACGGTGGCGCGGGGGGGCCCCCGGCCCCCCGGGTG 5217 Qy 609SerSerGlyAlaValGlyValMetPheHisAlaValproclyLySGluGlyAsp 626 .	Db 5218 ATCGACCCGAGCAACGGTGTCGAACTCAGCGGCACCGCCGAGCGTCCGGGTGATC 5277	Qy 627 LeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrp 645	Qy 646 SerAlalleThrGlyValSerSerAlaValAsnValGlyPheGlyLys 661	Db 5338 AGCTTCACGCCGGCGCCGCTGGCCAACGGCACGGTGGTCAATGCCGTGGCCCAGGAC 5397		Qy 682 GlyalaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHis 701	Db 5458	OY 702 GlnTyrGlyAsnTrpGlyGlnAlalleThrGlyAspHisAlaAsnLeuArgArgValTyr 721	

466 SerAlabudiyAspleuGiyGiyPherPhilis	5398 CCGGCCGCCATACCAGCGCCCGCCAGCACCAGCGGCGCCCACG 5457 682 GlyAlaTyrArgSerAspAspCysGlyThrTrpValLeulleAsnAspAspGlnHis 701 5458
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	8 8 8 8 8 8
10	373 ValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaileAspFroPheAsnSer 392 4477 GTG
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	6 6 6 6 6 6 6

Search completed: May 12, 2004, 18:05:33 Job time : 11480.2 secs

us-09-917-376-3.p2n.rnpb

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

May 11, 2004, 11:51:45; Search time 4.17339 Seconds (without alignments) 2051.340 Million cell updates/sec Run on:

US-09-917-376-4

470 1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX 89 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283366 seqs, 96191526 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	1,4-beta-glucanase	xylanase - Caldice	о О	(EC 3.	~	mannan endo-1,4-be	-	(EC 3	(EC 3	-	cellulase homolog	bifunctional cellu	cellulase (EC 3.2.	endoqlucanase I (E	cipa protein - Clo	endoglucanase - Er	cellulase (EC 3.2.	thermoactive cellu	scaffolding protei	probably celluloso	scaffolding protei	cellulose 1,4-beta	cellulose-binding	hypothetical prote	_	_	probable RTX famil	cal	d)	
ΩI	T31337	108	95	587	JN0111	A48954	T17120	A27198	A43802	802711	A41897	I40548	B41897	A47704	836859	839962	S54744	S12021	PC6006	C97012	T30433	847466	A44140	T08685	7	22	54	F90696	919	
DB	2	~	7	~	N	7	7	~	~	~	~	~	~	~	7	7	7	7	7	N	N	~	~	~	N	~	~	~	7	
Length	1711	7	0	508	499	1331	1742	499	915	1039	145	486	700	879	1854	505	504	986	586	1483	1162	1230	1848	618	547	42	5188	5291	574	
% Query Match	40.6	40.5		40.4	ö	e,		ę.	œ.	œ.	φ,	38.3	ġ.	'n,	'n.	i,	ı.	٥.	۲.	۲.	ė.	4.	4.	Ġ.		'n.			•	
Score		190.5	190	190	188	186.5	186.5	186	182.5	182.5	180.5	180	172	165.5		4	4	141.5	130.5	130.5	126.5	117	114	75.5	73.5	73.5	70	70	68.5	
Result No.		7	М	4	Ŋ	9	7	œ	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	

30 67 14.3 271 2 H72684 31 66.5 14.1 247 2 UC7201 33 65.5 13.9 751 2 T47790 34 65.5 13.9 1435 2 T47790 35 65 13.8 523 2 C70717 36 4 13.6 1571 2 S50669 38 64 13.6 1571 2 S50669 39 63.5 13.5 398 2 H70393 40 63.5 13.5 269 2 H95976 41 62.5 13.3 269 2 H95976 42 62.5 13.3 4936 2 H82512 44 62.5 13.3 4936 2 H82515 45 62 13.3 4936 2 H82515	hypothetical prote	thaumatin-like pro	hypothetical prote	ser-lys rich hypot	polyprotein - hepa	probable purH prot	probable reverse g	hypothetical prote	hensin - rabbit	hemolysin - Aquife	sanko - human	endo-beta-1,3-1,4-	replicative DNA he	protein-tyrosine-p	hypothetical prote	Ig heavy chain V r
	H72684	JC7201	T47790	T40462	T01075	C70717	G72614	850669	T30549	H70393	A59386	H95976	AB2512	B44390	AH2515	E37267
	7	~	N	7	~	7	7	7	~	7	7	7	7	ч	7	7
	271	247	535	751	1435	523	1222	1571	1594	398	2403	269	454	802	4936	122
33 31 32 33 33 34 36 65 65 65 65 65 65 65 65 64 40 65 65 65 65 65 65 65 65 65 65	14.3	14.1	14.0	13.9	13.9	13.8	13.6	13.6	13.6	13.5	13.5	13.3	13.3	13.3	13.3	13.2
	63	66.5	99	65.5	65.5	65	64	64	64	63.5	63.5	62.5	62.5	62.5	62.5	62
	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

RESULT 1
T31337
1,4-beta-glucanase (EC 3.2.1) - Anaerocellum thermophilum (fragment)
C;Species: Anaerocellum thermophilum
C.Date: 02-Sep-2000 #sequence revision 02-Sep-2000 #text change 02-Sep-2000
C;Accession: T31337
R; Zverlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K.
Microbiology 144, 457-465, 1998
A, Title: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from
omains.
A; Reference number: Z21003; MUID: 98154434; PMID: 9493383
A; Accession: T31337
A;Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A;Residues: 1-1711 <zve></zve>
A; Cross-references: EMBL: Z86105; NID: e1071329; PID: e350354; PIDN: CAB06786.1
C;Genetics:
A; Gene: celA
O Warmand and a large of bright of the

C; Keywords: glycosidase; hydrolase

Gaps 5 Length 1711; Query Match
40.6%; Score 191; DB 2; Length 17
Best Local Similarity 43.8%; Pred. No. 1.2e-11;
Matches 39; Conservative 18; Mismatches 30; Indels

2

1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59 a à

742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88 ઠ g

xylanase - Caldicellulosiruptor sp.

C;Species: Caldicellulosiruptor sp.

C;Species: Caldicellulosiruptor sp.

C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C;Accession: T11085

R;Morrist, D.D.; Gibbs, M.D.; Ford, M.; Thomas, J.; Bergquist, P.L.

submitted to the EMBL Data Library, December 1997

A;Description: Family 10 and 11 xylanase genes from Caldicellulosiruptor sp. Rt69B.1.

A;Reference number: 220972

A;Accession: T31085
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1779 < MOR>
A;Cross-references: EMBL:AF036924; NID:g2760905; PID:g2760909; PIDN:AAB95326.1
C;Genetics:

Query Match

Length 1779; DB 2; 40.5%; Score 190.5; G.

Gaps

4,

26; Indels

Length 508;

```
A.Residues: 10-13,'V',15,'N',17-18,'V',20-21,'F',23,'A',25-26,'AI',29-31,'P',33,'PQ',36-3A;Cross-references: GB:M38634; NID:g142657; PIDN:AAA22300.1; PID:g142658
A.Experimental source: strain ATCC 6633
C.Comment: The low molecular weight of the mature protein suggests carboxyl-terminal proc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Experimental source: strain DLG
A, Note: the authors believe Met-1 and Met-2 may be alternate initiators
C, Comment: The low molecular weight of the mature protein suggests carboxyl-terminal proc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Pathway: cellulose degradation
C;Keywords: alternative initiators; extracellular protein; glycosidase; hydrolase; polyst
F;1-38/Domain: (or 2-38) signal sequence #status predicted <SIG>
        A;Molecule type: DNA
A;Residues: 10-508 <WOL>
A;Residues: 10-508 <WOL>
A;Cross-references: EMBL:229076; NID:9509266; PIDN:CAA82317.1; PID:9509267
A;Cross-references: EMBL:229076; NID:9509266; PIDN:CAA82317.1; PID:9509267
B;Seo, Y.S.; Lee, Y.H.; Pek, U.H.; Kang, H.
Korean J. Microbiol. 24, 236-242, 1986
A;Title: Analysis on the nucleotide sequence of the signal region of Bacillus subitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A, Description: hydrolysis of 1, 4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-38/Domain: signal sequence #status predicted <SIG>
F;39-508/Product: cellulase #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    362 GVSVQYKAGDGRVNSNQIRPQLHIKNNGNATVDLKDVTARYWYNVKNKGQN---PDCDYA 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Species: Bacillus subtilis
Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Oct-1999
Accession: A26874; B26874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Residues: 1-508 <ROB1>
A;Cross-references: GB:M16185; NID:g143007; PIDN:AAA22496.1; PID:g143008
A;Experimental source: strain DLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacillus subtilis (strain DLG)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
40.4%; Score 190; DB 2; Length 508;
Best Local Similarity 43.0%; Pred. No. 4.2e-12;
Matches 37; Conservative 17; Mismatches 28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R,Robson, L.M.; Chambliss, G.H.
J. Bacteriol. 169, 2017-2025, 1987
A,Fitle: Endo-beta-1-4-Jucanase gene of Bacillus subtilis
A,Reference number: A26874; MUID:87194581; PMID:3106328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match

40.4%; Score 190; DB 2;
Best Local Similarity 40.7%; Pred. No. 4.2e-12;
Matches 35; Conservative 21; Mismatches 26;
                                                                                                                                                                                                                                                                                                                         A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         419 QIĞĞÜNVTHKFVTLHKPKQGADİYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N;Alternate names: endo-1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 AMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cellulase (EC 3.2.1.4) precursor
                                                                                                                                                                                                                                                A; Reference number: I39803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A, Accession: B26874
A, Molecule type: protein
A, Residues: 39-53 < ROB2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
                                                                                                                                                                                                                                                                                         A; Accession: I39803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: A2687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Description:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gene: bglC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gestinate names: endol.1, -beta-glucanase

Cilulase (EC 3.2.1.4) bglC precursor - Bacillus subtilis

NiAlternate names: endol.1, -beta-glucanase

C; ppecies: Bacillus subtilis

C; Date: 05-Dec-1997 #sequence revision 05-Dec-1997 #text change 21-Jul-2000

C; Accession: G69593; A26114; I40353; S44239; S49103; I39803

R; Kunst, F.: Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C; Bron, S.; Brounllet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.;

A; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A; Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauueell

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A; Authors: Schleich, S.; Schroeter, T.; Terpstra, P.; Tognoni, A.; Tosato, T.; Scron

Akouthi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Voshida, R.

A; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A; Reference number: A65580; MUID:98044033; PMID:9384377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A,Note: part of this sequence, including the amino end of the mature form, was confirmed R;Lindahl, V.; Aa, K.; Tronsmo, A. Antonie Van Leeuwenhoek 66, 327-332, 1984a-1,4-glucanase gene from Bacillus subtilis C A;Reference number: 140353; MUID:95225656; PMID:7710280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Experimental source: strain 168
R; MacKay, R.M.; Lo, A.; Willick, G.; Zuker, M.; Baird, S.; Dove, M.; Moranelli, F.; Seli Nucleic Acids Res. 14, 9159-9170, 1986
A; Title: Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.
A; Reference number: A26114; MUID:87066783; PMID:3024130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-508 <KUN>
;Cross-references: GB:299113; GB:AL009126; NID:g2634090; PIDN:CAB13696.1; PID:e1183471;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Description: Genes encoding beta glucan-hydrolyzing enzymes in Bacillus subtilis: cons
                                                                                                                                                                            4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Residues: 10-291,'N', 293-508 <LIN1>
A;Cross-references: EMEL:X67044; NID:g39776; PIDN:CAA47429.1; PID:g39777
R;Lindahl, V.; Aa, K.
Rubmitted to the EMBL Data Library, June 1992
A;Reference number: S24239
A;Accession: S24239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 10-291,'N', 293-508 <LIN2>
A; Cross-references: EMBL:X67044; NID:939776; PIDN:CAA47429.1; PID:939777
A; Experimental source: strain CK-2
B; Wolf, M.; Gecal, A.; Borriss, R.
submitted to the EMBL Data Library, December 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Residues: 10-508 <MAC>
A;Cross-references: GB:X04689; NID:g39823; PIDN:CAA28392.1; PID:g39824
A;Experimental source: strain PAP115
                                                    Indels
1.4e-11;
ches 33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Status: preliminary; translated from GB/EMBL/DDB.
        ; Pred. No. 1.4e
14; Mismatches
                                                                                                                                                                                                                                                                                                                                                         1172 IGASNVTFNFVKLTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                 64 MGCGNIRASFGSVNPATPTADTYLQ 88
            43.5%;
                                                37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Reference number: S49103
    Best Local Similarity
Matches 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Accession: I40353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: S49103
                                                                                                                            ઠ
                                                                                                                                                                                        원
                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                     셤
```

DIG.

62

Gaps

4.

61

셤

```
cellulase (EC 3.2.1.-) precursor, thermoactive - Caldocellum saccharolyticum C;Species: Caldocellum saccharolyticum C;Species: Caldocellum saccharolyticum C;Species: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000 C;Accession: T17120; A4374.
R;Te'o, V.S.; Saul, D.J.; Bergquist, P.L.
Appl. Microbiol. Biotechnol. 43, 291-296, 1995
A;Title: CelA, another gene coding for a multidomain cellulase from the extreme thermophi A;Reference number: Z18698; MUID:95336703; PMID:7612247
A;Accession: T17120
A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A.Residues: 1-1742 <TEO>
A.Cross-references: EMBL:L32742; NID:g537499; PID:g537500; PIDN:AAA91086.1
R.Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl. Environ. Microbiol. 57, 694-700, 1991
A.Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding A;Reference number: A43745; MUID:91247819; PMID:2039230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A,Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A,Pathway: cellulose degradation C,Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation F;1-36/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  704 SGQIKVLYANKETNSTINTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAIS-DW 762
                                                                                              364 SGQIKVLYANKETNSTINTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAIS-DW 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Cipecies: Bacillus subtilis
Cipate: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 15-Oct-1999
Cipate: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 15-Oct-1999
Cipacession: A27198
Rivakamura, A. Pocumi, T.; Beppu, T.
Bur. J. Biochem. 164, 317-320, 1987
A; Title: Nucleotide sequence of a cellulase gene of Bacillus subtilis.
A; Accession: A27198
A; MulD: 87190397; PMID: 3106035
A; Accession: A27198
A; MulD: 87190397; PMID: 3106035
A; Cross-references: GB: M28332; NID: G142670; PIDN: AAA22307.1; PID: G142671
A; Experimental source: strain IF03034
C; Function:
                                SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Scatus: preliminary
A;Molecule type: DDA
A;Residues: 1516-1544, A',1546-1742 <LUB>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72860.1; PID:g144293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain IFO3034) N, Alternate names: endo-1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1742;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      / Match 39.7%; Score 186.5; DB 2. Local Similarity 42.5%; Pred. No. 3.7e-11; nes 37; Conservative 18; Mismatches 31.
                                                                                                                                                                                                                       423 AQIGASNVTFKFVKLSSSVSGADYYLE 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                           88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                           AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A,Gene: celA
C,Keywords: glycosidase; hydrolase
                                    N
                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                  셤
                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                           ð
                                    Š
                                                                                                                                                                                                 Cellulaee (EC 3.2.1.4) precursor - Bacillus subtilis (strain BSE616)
N.Alternate names: endo-1,4-beta-glucanase
C.Species: Bacillus subtilis
C.Species: Bacillus subtilis
C.Species: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 28-May-1999
C.SAccession: UN0111
R.Park, S.H.; Kim, H.K.; Pack, M.Y.
Agric. Biol. Chem. 55, 441-448, 1991
A.Title: Characterization and structure of the cellulase gene of Bacillus subtilis BSE61
A.Reference number: UN0111; MUID:91299280; PMID:1368694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce A;Pathway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;1-29/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N;Alternate names: beta-mannanase
C;Species: Caldocellum saccharolyticum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A48954; B43745
R;Gibbs, M.D.; Saul, D.J.; Luthi, E.; Bergquist, P.L.
Appl. Environ. Microbiol. 58, 3864-3867, 1992
A;Title: The beta-mannanase from "Caldocellum saccharolyticum" is part of a multidomain A;Reference number: A48954; MUID:93119139; PMID:1476429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene coding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Cross-references: GB:D01057; NID:g216387; PIDN:BAA00859.1; PID:d1001323; PID:g216388 A;Note: the authors translated the codon ATA for residue 102 as Tyr C;Comment: The low molecular weight of the mature protein suggests carboxyl-terminal pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Modecule type: nucleic acid
A; Residues: 1-131 (GIBs)
A; Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A; Cross-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A; Note: sequence extracted from NCBI backbone (NCBIN:121576, NCBIP:121577)
R; Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl: Environ. Microblol. 57, 694-700, 1991
A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a A; Reference number: A41745; MUID:91247819; PMID:2039230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        mannan endo-1,4-beta-mannosidase (EC 3.2.1.78) - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Residues: 1-337,'PPROHOHRO' <LUE>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72861.1; PID:g144294
A;Note: the authors translated the codon CAC for residue 262 as Glu
A;Note: this sequence has been revised in reference A48954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40.0%; Score 188; DB 2; Length 499; 40.0%; Pred. No. 6.7e-12; ative 18; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 39.7%; Score 186.5; DB 2; Length Best Local Similarity 42.5%; Pred. No. 2.8e-11; Matches 37; Conservative 18; Mismatches 31; Indels
|||||: | :: |||||: QMGCGNLTHKFVTLHKPKQGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   :||||:
411 LGCGNVTYKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 MGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 40.0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues: 1-499 < PAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: B43745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: JN0111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Molecule type:
```

吊 ò g

8

61

Gaps

ä

```
F;29-1039/Product: cellulase #status predicted <MAT> F;72-373/Domain: Streptomyces endo-1,4-beta-xylanase A homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Residues: 1-145 < HAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Residues: 1-486 <RES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Accession: I40548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Gene: cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C;Species: Caldoceium control of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of Sepisor of 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :Title: Nucleotide sequence of a gene from Caldocellum saccharolyticum encoding for exc
Reference number: $02711; MUID:89098398; PMID:2789517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cellulase (EC 3.2.1.4) / cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91) - Caldocellum Nallernate namese: endo-1,4-beta-glucanase C;Species: Caldocellum saccharolyticum (C;Species: Caldocellum saccharolyticum (C;Date: 30-Jan-1993 #sequence_revision 30-Sep-1993 #text_change 10-Jul-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Pathway: cellulose degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A homology
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-28/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as Pathway: cellulose degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N;Alternate names: endo_1,4-beta-glucanase
N;Contains: cellulase (EC 3.2.1.4); cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91)
C;Species: Caldocellum saccharolyticum
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C;Accession: S02711
                                                                                     2;
                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: DNA
A;Residues: 1-1039 <SAU>
A;Cross-references: EMBL:X13602; NID:g40645; PIDN:CAA31936.1; PID:g40646
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Pathway: cellulose degradation
C; Superfamily: Streptomyces endo-1,4-beta-xylanase A homology
C; Keywords: glycosidase; hydrolase; polysaccharide degradation
F;20-320/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>
                                                                                         4 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 2; Length 915;
39.6%; Score 186; DB 2; Length 499; ilarity 40.7%; Pred. No. 1.1e-11; Conservative 19; Mismatches 28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cellulase (EC 3.2.1.4) precursor - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38.8%; Score 182.5; DB 2
41.9%; Pred. No. 4.8e-11;
tive 18; Mismatches 31
                                                                                                                                                                                                                                                                                                                                                                                       :||||:
410 QIGCGNLTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QIGASNVTFKFVKLSSSVSGADYYLE 433
                                                                                                                                                                                                                                                                                                                                    AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: EMBL:X13602
C;Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 41.99
Matches 36, Conservative
                                              Local Similarity
hes 35; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Molecule type: DNA
Residues: 1-915 <SAU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Accession: S02711
                                                                                                                                                                                                                                                                                                                                        63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
        Query Match
                                              Best Loc
Matches
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
```

```
cellulase homolog - Bacillus lautus (fragment)
C;Species: Bacillus lautus
C;Species: Bacillus lautus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 15-Oct-1999
C;Accession: A41897; S2498
C;Accession: A41897; S2498
A;Bacteriol. 174, 3522-3531, 1992
A;Title: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-1,4,7,8eference number: A41897; MUID:92276330; PMID:1592807
A;Accession: A41897
A;Adolecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Accession: I40548
K;Han, S.J.; Yoo, Y.J.; Kang, H.S.
J. Biol. Chem. 270, 26012-26019, 1995
A;Title: Characterization of a bifunctional cellulase and its structural gene: the cel generence number: 140548; WUID:96029707; PMID:7592793
                                                                                                                                                                                                                                                                                                                                                  62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | : | : | : | : | | | | | : | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: Bacillus sp.
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7 VOYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                      3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: EMBL:M76588; NID:g142661; PIDN:AAA22302.1; PID:g142662
A;Experimental source: PL236
A;Note: sequence extracted from NCBI backbone (NCBIP:104604)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A, Cross-references: EMBL:U27084; NID:g857575; PIDN:AAC43478.1; PID:g857576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                          ij
      Length 1039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        38.3%; Score 180; DB 2; Length 486; 38.8%; Pred. No. 4.5e-11; ive 18; Mismatches 32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match

38.4%; Score 180.5; DB 2; Length
Best Local Similarity 43.9%; Pred. No. 1.1e-11;
Matches 36; Conservative 18; Mismatches 25; Indels
Query Match 38.8%; Score 182.5; DB 2; Best Local Similarity 41.9%; Pred. No. 5.5e-11; Matches 36; Conservative 18; Mismatches 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bifunctional cellulase precursor - Bacillus sp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :| |: | :: | |: | |: | 179 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     67 GNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  33; Conservative
```

g

8

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1615,1617-1854 <GE2>
A;Residues: 1-1615,1617-1854 <GE2>
A;Cross-references: EMBL:L08665
R;Fujino, T.; Beguin, P.; Aubert, J.P.
FEMS Microbiol. Lett. 94, 165-170, 1992
A;Title: Cloning of a Clostridium thermocellum DNA fragment encoding polypeptides that bi A;Reference number: S25767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A, Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulosom?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: EMBL:L08665
R; Gerngross, U.T.; Romaniec, M.P.M.; Kobayashi, T.; Huskisson, N.S.; Demain, A.L.
Mol. Microbiol. 8, 325-334, 1993
A; Title: Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal
A; Reference number: S33527; WUID:93302508; PMID:8316083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Accession: $28659
A;Molecule type: DNA
A;Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1768,'R',1770-1854
A;Cross-references: EMBL:X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: DNA
A;Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1854 <FUJ>
A;Cross-references: EMBL:X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                   01-Dec-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       365 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:X67506; NID:g296879; PID:g296880; PIDN:CAA47840.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       35.1%; Score 165; DB 2; Length 1854; 36.1%; Pred. No. 6.9e-09; tive 23; Mismatches 29; Indels 10
                                     cipA protein - Clostridium thermocellum N;Alternate names: probable cellulosome protein large chain SL c;Species: Clostridium thermocellum C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change C;Accession: 836659; 833527; $25767; $28659; T18261 R;Gerngross, U.T.; Demain, A.L. submitted to the EMBL Data Library, January 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Reference number: Z18847; MUID: 93209931; PMID: 8458832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1821-1854 <FU2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R;Beguin, P. submitted to the EMBL Data Library, August 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R.Fujino, T.; Beguin, P.; Aubert, J.P.
J. Bacteriol. 175, 1891-1899, 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                 A;Reference number: S36859
A;Accession: S36859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Reference number: S28659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1854 <GER>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Accession: T18261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A, Accession: S25767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: cipA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Species: Clostridium thermocellum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C;Accession: A7704
B;Hazlewood, G;P.; Davidson, K; Laurie, J.I.; Huskisson, N.S.; Gilbert, H.J.
J; Gen. Microbiol. 139, 307-316, 1993
A;Title: Gene sequence and properties of Cell, a family E endoglucanase from Clostridium
A;Reference number: A47704; MUID:93171873; PMID:8436949
                                                                                                                                                                                                                                                                                                                                                                                 N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus lautus
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Oct-1999
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Oct-1999
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Oct-1999
R;Acternion: 174, 3522-3531, 1992
A;Fitle: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-1,4
A;Reference number: A41897; MUID:92276330; PMID:1592807
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A; Pathway: cellulose degradation
C; Keywords: glycosidase; hydrolase; polysaccharide degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;Residues: 1-700 <HAN>
;Cross-references: EMBL:M76588; NID:g142661; PIDN:AAA22303.1; PID:g142663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: GB:L04735; NID:g144807; PIDN:AAA20892.1; PID:g144808
A;Note: sequence extracted from NCBI backbone (NCBIN:125637, NCBIP:125638)
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36.6%; Score 172; DB 2; Length 700; 44.9%; Pred. No. 4.5e-10; tive 18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      endoglucanase I (EC 3.2.1.-) Cell - Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Experimental source: PL236
A;Note: sequence extracted from NCBI backbone (NCBIP:104605)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ||| :| || :| || : | | | : | | | : | | | : | | | : | | | : | | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                 LGCGNVSHTVVTLHKPKQGADTYLE 435
MGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                           cellulase (EC 3.2.1.4) - Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 44.9%
nes 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Status: preliminary
Molecule type: DNA; protein
Residues: 1-879 <HAZ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Molecule type: DNA
     64
                                                                                                       411
```

Function

Matches

셤 8 g

Š

Search completed: May 11, 2004, 12:10:26 Job time : 6.17339 secs

61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88

셤 8 원

ï

Gaps

1;

9

g δ

ઠે

RESULT 15

1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD

35.2%; Score 165.5; DB 2; Length 879; 33.0%; Pred. No. 2.7e-09; ... tive 26; Mismatches 32; Indels 1

Conservative

Query Match Best Local Similarity

```
P07983 bacillus su
P10475 bacillus su
P25549 bacillus su
P22534 caldocellum
P10474 cendogluca
P29719 paenibacill
P29719 paenibacill
P29719 paenibacill
P29719 paenibacill
P29719 paenibacill
P29719 paenibacill
P29719 erwinia car
Q59394 erwinia car
Q59395 erwinia car
P50900 clostridium
P23659 clostridium
P23659 clostridium
P23659 clostridium
P23659 clostridium
P2365 clostridium
P2365 clostridium
P2365 clostridium
P2365 clostridium
P2365 clostridium
P2365 clostridium
P2365 clostridium
P2366 saccharomyc
P7719 thermoplasm
Q9469 homo sapien
Q9469 homo sapien
Q9465 mus musculu
Q9955 mus musculu
Q9955 mus musculu
Q9955 mus musculu
Q9955 mus musculu
Q99515 pseudomonas
P50937 rhodobacter
P3081 pseudomonas
                                                                                                                                                                                   May 11, 2004, 11:41:35 ; Search time 2.42639 Seconds (without alignments) 1909.933 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                         470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
   5:1.6
Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141681 seqs, 52070155 residues
GenCore version
Copyright (c) 1993 - 2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUN1 BACSU
GUN2 BACSU
GUN2 BACSU
GUN3 CALSA
GUN3 CALSA
GUN3 CALSA
GUN4 CALSA
GUNA CALSA
GUNA PAELA
GUN7 CLOTM
CIPA CLOTM
CIPA CLOTM
GUNV ERWCA
GUNV ERWCA
GUNV ERWCA
GUNY ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUNZ_CLOSR
CGBA_CLCCL
HVJ9_MOUSE
FLA1 METVA
TP1A_MALDO
PUS9_MYCTU
ILVD THETVA
ATCS_YEAST
FLA1 THEVO
TY3H SCHMA
OLO6_HUMAN
EXOM_RHIME
EXOM_RHIME
STL2_MOUSE
STL2_MOUSE
STL2_MOUSE
HIGE PROSERL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                          - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                           US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SwissProt 42:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             444

986

1086

1188

1118

1129

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

1200

12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        %
Query
Match I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          190
186.5
186.5
186.5
182.5
10.5
165.5
165.5
149.5
147.5
147.5
147.5
141.5
141.5
141.5
141.5
141.5
141.5
141.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165.5
165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64.5
64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63.5
63.5
63.5
63.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62.5
62.5
                                                                                                                                                                                                                                                                                                                           Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65
                                                                                                                                 OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                                                                                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        333338432777777777779
```

61.5 13.1 328 1 HAIQ_MOUSE 61.5 13.1 489 1 GATA_CAUCR 61.5 13.1 678 1 YIHQ_ECOLI 61 13.0 120 1 YYHQ_ECOLI 61 13.0 338 1 RTCA_ECOL6 61 13.0 338 1 RTCA_ECOL6 61 13.0 342 1 RTCA_ECOL6 60.5 12.9 1157 1 C9CA_BACTO 60 12.8 1226 1 PAT2_CAEEL 59.5 12.7 261 1 COTN_BACSU	ALIGNMENTS	SU 1_BACSU STANDARD; PRT; 499 AA.	P0/301; 01-AUG-1988 (Rel. 08, Created) 01-NOV-1991 (Rel. 20, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update) Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)	(Cellulase). BGLC OR GLD. Bacilus subcilis. Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillu NCBI_TaxID=1423;	 PubMed=3106328; bliss G.H.;	"Endo-beta-1, 4-glucanase gene of Bacillus subfills DLG."; J. Bacteriol. 169:2017-2025 (1987) -!- CATALYTIC ACTUTIY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans!- SIMILARITY: Belongs to cellulase family A (family 5 of glychydrolases).		Bend an email to license@isb-sib.ch). L; Mi6185; AAA22496.1; ALT_INIT. ?, A26874; A26874. P; 085465; 1A3H.	5 . P5, 1.	Cellulose degradation; Hydrolase; Glycosidase; Signal. 2 SIGNAL. 2 49 ENDOGLUCANASE. ACT_SITE 169 169 PROTON DONOR (BY SIMILARITY). ACT_SITE 257 257 NUCLEOPHILE (BY SIMILARITY). DOMAIN 350 499 AM: 339D04EB55A63EBI CRC64:	ttch 40,4%; Score 190; al Similarity 43.0%; Pred. No. 1
₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩		20.55	AC F0/9 DT 01-A DT 28-F DE Endo	DE (Cel GN BGLC OS Baci OC Bact			CC This CC Detw CC the CC use CC modi		DR Inte DR Inte DR Inte DR Pfam DR Pfam DR Prob	FT SIGNAL FT CHAIN FT ACT SI FT ACT SI FT ACT SI FT DOMAIN	Query Match Best Local

GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA 62

ઠે

```
ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kunst F., Ogaswara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriss F., Brans A., Braun M., Brignell S.C., Bron S., Brouilet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F. Y., Devinet K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Britz C., Pujita M., Fabret C., Ferrari E., Foulger D., Fritz C., Pujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Asahara Y., Klaerr-Blanchard M., Klein C., Kurita R., Lapidus A., Lardinois S., Lauber J., Madigue C., Mediau N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parsecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Resecan E., Pujic P., Purnelle B., Rapoport G., Roy M., Sanlan E., Schleich S., Schroeter R., Scoffone F., Sanlan E., Schleich S., Serror R., Stror P., Shin B.S., Soldo B., Sorokin A., Takemaru K., Jakemaru K., Takemaru K., Sorokin A., Takemaru K., Sorokin A., Takemaru K., Takemaru K., Takemaru K.,
Lindahl V., Aa K., Tronsmo A.;
"Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIINE=97124194; PubMed=8969507;
Rose M., Entian K.D.;
"New genes in the 170 degrees region of the Bacillus subtilis genome
encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino
                                                                                                                                                                                                                                                                                                                                    Mackay R.M., Lo A., Willick G., Zuker M., Baird S., Dove M., Moranelli F., Seligy V.; "Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene."; Nucleic Acids Res. 14:9159-9170(1986).
                                                                                                                                                          01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
10-CTT-2003 (Rel. 42, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Carboxymethyl-cellulase) (CWCASE) (Cellulase).
EGLC OR GLD OR EGLS OR BSU18130.
                                                                                                                                                                                                                                                       Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus
                                                                                                                                     499 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Antonie Van Leeuwenhoek 66:327-332(1994).
                                                           410 QMGCGNLTHKFVTLHKPKQGADTYLE 435
                                         63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                        MEDLINE=87066783; PubMed=3024130;
                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=95225656; PubMed=7710280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=98044033; PubMed=9384377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Microbiology 142:3097-3101(1996).
                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                         [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                          Bacillus subtilis.
                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             transporter.
                                                                                                                                                                                                                                                                  NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                          STRAIN=PAP115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       subtilis CK-2.
                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=CK-2;
                                                                                                                                   3UN2 BACSU
P10475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=168;
                                                                                                        RESULT 2
GUNZ_BACSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             acid
               엄
                                        ઠે
                                                               g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lebib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
Winners P., Wippt A., Yamanoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
"The complete genome sequence of the Gram-positive bacterium Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                       AREALNE-S2225655; PubMed=7710279;
AREDLINE-95225655; PubMed=7710279;
AR K., Flengsrud R., Lindahl V., Tronsmo A.;
Characterization of production and enzyme properties of an endo-beta-1,4-glucanase from Bacillus subtilis CK-2 isolated from compost soil.";
Antonie Van Leeuwenhoek 66:319-326(1994).
-!- CATALYTITY: Endobydolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucons.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENDOGLUCANASE.
PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
S -> N (IN REF. 2).
9 8F735FF711B3EAEZ CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 40.4%; Score 190; DB 1; Length 499; Best Local Similarity 40.7%; Pred. No. 1.3e-13; Matches 35; Conservative 21; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Probom; PD001947; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :||||:
410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; Z29076; CAA82317.1; --
EMBL; X04689; CAA28392.1; --
EMBL; X7044; CAA47429.1; --
EMBL; Z7234; CAA97610.1; ALT_INIT.
EMBL; Z99113; CAB13696.1; ALT_INIT.
EMBL; Z99113; CAB13696.1; ALT_INIT.
HSSP, 068465; 1A3H.
Subtilist; BG10437; bg1C.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
Pfam; PF00150; Cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29 EB
169 PP
169 PP
169 CB
283 CS
55287 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUN3 BACSU STANDARD;
P23549;
01-NOV-1991 (Rel. 20, Created)
                                                                                                                                                                                                    Nature 390:249-256(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           169
257
350
283
283
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Complete proteome.
                                                                                                                                                                                                                                                            SEQUENCE OF 30-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        hydrolases)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                         subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 3
GUN3_BACSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

m

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNCDW
Caldicellulosiruptor
          NCBI_TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
 ઠે
                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MANA.
Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P22533;
01-AUG-1991 (Rel. 19, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-Beta-mannanase) (Endo-1,4-
                                                                                                                                                                         Agric. Biol. Chem. 55:441-448(1991).
-!- CATALVITIC ACTIVITY: Enchydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                          Park S.H., Kim H.K., Pack M.Y.; "Characterization and structure of the cellulase gene of Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULOSE-BINDING (BY SIMILARITY)
2E821E3D8BBACA04 CRC64;
 01-NOV-1991 (Rel. 20, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Carboxymethyl-cellulase) (CMCASE) (Cellulase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40.0%; Score 188; DB 1; Length 499; 40.0%; Pred. No. 2.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROTON DONOR (BY SIMILARITY) . NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34; Conservative 18; Mismatches 31; Indels
                                                                    Bacteria, Firmicutes, Bacillales, Bacillaceae, Bacillus
NCBI_TaxID=1423,
                                                                                                                                                                                                                                                                                                                                                                        InterPro; PR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001547; Glyco_Dydro_5.
Pfam; PF00150; CBllulase; 1.
Probom; PD00150; CBllulase; 1.
Probom; P000150; GlYCOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1331 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENDOGLUCANASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :||||:
411 LGCGNVTYKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  64 MGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                   STRAIN=BSE616;
MEDLINE=91299280; PubMed=1368694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55169 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40.08;
                                                                                                                                                                                                                                                                                                                                         EMBL; D01057; BAA00859.1; -. PIR; JN0111; JN0111.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              169
257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             169
257
257
350
499 AA;
                                                                                                                                                                  subtilis BSE616.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                  HSSP; 085465; 1A3H
                                                         Bacillus subtilis
                                                                                                                                                                                                                         hydrolases).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MANB CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY A (FAMILY S OF GLYCOSYL HYDROLASES).
-!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY J (FAMILY 44 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                    galacroglucomannans.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- MISCELLANEOUS: This enzyme is most active at pH 6 and 80 degrees
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                      MEDLINE=93119139; PubMed=1476429;
Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.;
"The beta-mannanase from 'Caldocellum saccharolyticum' is part of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
BETA-MAUNANSE/ENDOGLUCANASE A.
GATALYTIC (MAUNANASE ACTIVITY).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
PRO/SER/THR-RICH (PT BOX).
CATALYTIC (ENDOGLUCANASE ACTIVITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 1331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     146892 MW, FFBCAS1BBBDBF0E0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTON DONOR (BY SIMILARITY) NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProDom; PD001947; CBD 3; 2.
PROSITE; PS00659; GLYCOSYL HYDROL F5; 1.
Hydrolase; Glycosidase; Cellulose degradation; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         / match
Local Similarity 42.5%; Pred. No. 1e-12;
les 37; Conservative 18; Mismatches 2
                                                                                                                              Appl. Environ. Microbiol. 58:3864-3867(1992).
[2]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              T -> P (IN REF.
TPTPTPT -> ROHO!
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A48954; A48954.
HSSP; Q06851; INBC.
INTEPPO; IPR001956; CBD 3.
INTEPPO; IPR001956; CelTul bind.
InterPro; IPR001547; Glyco_hydro_5.
Pfam; PF00942; CBM 3; 2.
Pfam; PP00150; celTulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; L01257; AAA71887.1; -. EMBL; M36063; AAA72861.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Multifunctional enzyme
                                                                                                               multidomain enzyme."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1331 AA;
FROM N.A.
```

```
SIGNAL
                                                                                                   DOMAIN
                                                                                                                     DOMAIN
                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GUNB_CALSA
                                                                                                                                                                                                                                                                                                                                       Best Loc
Matches
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY B (FAMILY 9 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULOSE.
-!- CAPALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- PTM: The linker region (also termed "hinge") may be a potential , site for proteolysis.
                                                                                                                                                                                                                                                                                                 Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οţ
                                                                                                                                                                                                  01-AUG-1991 (Rel. 19, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase A)
(Cellulase A).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=91247819; PubMed=2039330;
Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
"Cloning, sequence analysis, and expression in Escherichia coli of
gene coding for a beta-mannanae from the extremely thermophilic
bacterium 'Caldocellum saccharolyticum'.";

-!- FUNCTION: THE N-TERMINAL DOWAIN OF CELA ENCODES FOR AN
ENDOGLUCANASE ACTIVITY ON CARBOXYMETHYLCELULOSE. THE C-TERMIN
DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINES-9536703; PubMed=7612247;
Te'O V.S., Saul D.J., Bergquist P.L.;
"celA, another gene coding for a multidomain cellulase from the extreme thermophile Caldocellum saccharolyticum.";
Appl. Microbiol. Biotechnol. 43:291-296(1995).
                                                                                                                                                                     PRT; 1742 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L (FAMILY 48 OF GLYCOSYL HYDROLASES).
                                                                        423 AQIGASNVTFKFVKLSSSVSGADYYLE 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP, P26221; 1TF4.
INTERPO. IPRO01956; CBD 3.
INTERPO: IPRO00855; Cellul bind.
INTERPO: IPRO00856; Glyco_hydro_48.
INTERPO: IPRO00856; Glyco_hydro_9.
INTERPO: IPRO00828; Glyco_trans_6hp.
Pfam; PF00942; CBM 3; 3.
Pfam; PF00719; Glyco_hydro_48; 1.
PRINTS; PR000844; GLHYDRAASE48.
Probom; PD001947; CBD_3; 2.
                                                   62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M36063; AAA72860.1; -.
BMBL; L01257; -; NOT ANNOTATED_CDS
PIR; T17120; T17120.
HSSP; P26221; 1TF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1516-1742 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; L32742; AAA91086.1; -.
                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=44001;
                                                                                                                                                                     CALSA
                                                   8
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose and celloteraose, releasing cellobiose from the non-reducing ends of the chains.
-!- SIMILARITY: In the N-terminal section; belongs to cellulase family F (family 10 of glycosyl hydrolases).
-!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
28-FEB-2003 (Rel. 11, Last annotation update)
Endoglucanase/exoglucanase B precursor [Includes: Endoglucanase
(EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
(Cellobiohydrolase); Exoglucanase (EC 3.2.1.91) (Exocellobiohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Saul D.J., Williams L.C., Love D.R., Chamley I.W., Bergquist P.I.;
"Nucleotide sequence of a gene from Caldocellum saccharolyticum
encoding for exocellulase and endocellulase activity.";
Nucleotic Acids Res. 17:439-439(1989).
-!- FUNCTION: THIS PROTEIN IS MADE UP OF TWO DOMAINS: THE N-TERMINAL
DOMAIN HAS EXOGLUCANASE ACTIVITY WHILE THE C-TERMINAL DOMAIN IS
                                                                                                                                                                                                LINKER ("HINDE") (PRO-THR BOX).
CELLUIOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
CATALYTIC 2.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
T -> A (IN REF. 2).
T -> A (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Η;
                                                                                                                                                             CELLULOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
                                                                                                                                                                                                                                                                                                                                                                                                             Length 1742;
ProDom; PD011903; Glyco hydro 48; 1.
PROSTIE: PS00592; GLYCOSYL HYDROL F9_2; 1.
PROSTIE: PS00698; GLYCOSYL HYDROL F9_2; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                          .5; b.
1.4e-12;
31;
                                                                                                                              CATALYTIC 1.
LINKER ("HINGE")
                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Pred. No. 1.4e-
18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 1039 AA
                                                                                                            ENDOGLUCANASE
                                                                                                                                                                                                                                                                                                                                                                                                                 Score 186.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           763 AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=89098398; PubMed=2789517;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Cellobiohydrolase); Exoglucane (1,4-beta-cellobiohydrolase)].
                                                                                                                                                                                                                                                                                                                                                                     193696
                                                                                                                                                                                                                                                                                                                                                                                                                                                       37; Conservative
                                                                                                                                                                                                                                                                                                                                                                       Ä,
                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI TaxID=44001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GUNB CALSA
P10474;
                                                                                                                                                                                                                                                                                                                             ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                        DOMAIN
```

Gaps

3,

Indels

Length 145;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                       7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC 66
                                                                                                                                                                                                                                                                                               endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- PTM: A SHORT FORM (EGA-S) ARISES FROM POSTTRANSLATIONAL PROTECLYSIS OF APPROXIMATELY 150 AA AT THE C-TERMINUS OF EGA-L.
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY J (FAMILY 44 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hansen C.K., Diderichsen B., Joergensen P.L.; "celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Paenibacillus lautus (Bacillus lautus).
Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-BPR-1993 (Rel. 41, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Cellulase A) (EG-A).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36.6%; Score 172; DB 1; Length 700;
                                                                                                                                              SEQUENCE 145 AA; 15782 MW; 9514E3A71B106AEB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; B41897; B41897.
HSSP; Q06851; UNBC.
InterPro; IPR001956; CelTul_bind.
Prom; PR00942; CBM 3; 1.
Prodom; PR001497; CBD 3; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3D5C8CADA53EEE0F CRC64;
                                                                                                                                                                                     ch 38.4%; Score 180.5; DB 1; Similarity 43.9%; Pred. No. 3.4e-13; 36; Conservative 18; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIÁL.
ENDOGLUCANASE A.
CELLULOSE-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          700 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                           67 GNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                   60 SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=PL236;
MEDLINE=92276330; PubMed=1592807;
HSSP, Q06851; 1NBC.
InterPro; 1PR001956; CBD 3.
InterPro; 1PR008965; Cellul_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 76910 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M76588; AAA22303.1; -.
                                               Pfam, PF00942; CBM 3; 1.
ProDom; PD001947; CBD 3; 1.
Hypothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               33
700
700
                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   548 7
213 2
700 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYDROLASES) .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GUNA PAELA
P29719;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
ACT SITE
SEQUENCE
                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                              KESULT 8
GUNA_PAELA
  S T X R R R F S
                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
  (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hansen C.K., Joergensen P.L., Diderichsen B.;
cabh from Bacillus lautus PL236 encodes a novel cellulose-binding
endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                         ENDOGLUCANASE/EXOGLUCANASE B.
THR/PRO-RICH, TANDEM REPEATS OF T-P.
CELLUJOSE-BINDING (BY SIMILARITY).
THR/PRO-RICH, TANDEM REPEATS OF T-P.
PROTON DONOR (POTENTIAL).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38.8%; Score 182.5; DB 1; Length 1039; 41.9%; Pred. No. 2.1e-12; 1tive 18; Mismatches 31; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1039 AA; 117641 MW; 0E0378171594DDAE CRC64;
                                                                                                                                                                                                        Pfam; PF00942; CBM 3; 1.
Pfam; PF00150; cellulase; 1.
Pfam; PF00150; cellulase; 1.
Prim; PF00131; dlycohydro.10; 1.
PRODOM; PD00134; dLyChyDRLASE10.
PRODITS; PR00134; dLyColl; 1.
SWART; SM00633; dlycoll; 1.
PROSITE; PS00659; GLYCOSYL HYDROL F10; 1.
Cellulose degradation; Hydrolase; Glycosidase; Repeat; Multifunctional enzyme; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-APR'1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-ARR-1993 (Rel. 40, Last annotation update)
Hypothetical protein in cela 5'region (Fragment)
Paenibacillus lautus (Bacillus lautus)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       145 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BY SIMILARITY
  entities requires a license agreement (Se
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :| |: | :: | |: 479 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                 HASSP, Q06651, INBC.
InterPro, IPR001956; CBD 3.
InterPro, IPR001006; CelTul bind.
InterPro, IPR001000; Glyco hydro 10.
InterPro, IPR001547; Glyco_hydro_5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN=PL236;
MEDLINE=92276330; PubMed=1592807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M76588; AAA22302.1; -.
                                                                EMBL; X13602; CAA31936.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                               1039
                                                                                      PIR; S02711; S02711.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A41897; A41897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
Les 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YCEA PAELA P29718;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
```

YCEA_PAELA

RESULT 7

Matches

셤 ò

```
CLOTM
                                                                                                                                                                                                                                                                                                                               CIPB_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CIPB.
                                                                                                                                                                                                                                                                                                       RESULT 10
                                                                                                                                                 g
                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                   HANDER REPRESENTATION OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE P
                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                         1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                               01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Gene sequence and properties of Cell, a family E endoglucanase from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Clostridium thermocellum.",
J. Gen. Microbiol. 139:307-316(1993).
-!- FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-GLUCANS. PRINCIPALLY ACTIVE AGAINST BARLEY BETA-GLUCAN.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- PATHWAY: Cellulose degradation.
-!- SIMILARITY: Belongs to cellulase family E (family 9 of glycosylhydrolases).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clostridium thermocellum.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
                          4,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=93171873; PubMed=8436949;
Hallewood G.P., Davidson K., Laurie J.I., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPRO1956; CBD 3.
InterPro; IPRO1956; CBD 3.
InterPro; IPRO1956; Cellul bind.
InterPro; IPRO1956; Cellul bind.
InterPro; IPRO1929; Glyco_fydro-9.
InterPro; IPRO1959; Glyco_trans_6hp.
Pfam; PF00942; CBM 3; 2.
Pfam; PF00942; CBM 3; 2.
Probom; PD001947; CBD 3; 1.
PROSITE; PS00592; GLYCOSYL HYDROL F9_1; 1.
PROSITE; PS00698; GLYCOSYL HYDROL F9_2; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
SIGNAL
  ; Pred. No. 2e-11;
18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35A60069A514A927 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CELLULOSE-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENDOGLUCANASE I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                         879 AA
                                                                                                                                                                            60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A., AND SEQUENCE OF 56-69
                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97796 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L04735; AAA20892.1; -.
44.98;
                      40; Conservative
                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          879
518
879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A47704; A47704.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       879 AA;
Best Local Similarity
Matches 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=NCIB 10682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P26221; 1TF4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gilbert H.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                      (Cellulase I)
                                                                                                                                                                                                                                                                                                         CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                    002934:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
DOMAIN
                                                                                                                                                                                                                                                            RESULT 9
GUNI_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                         ò
                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage, by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tidentification of the cellulose-binding domain of the cellulosome subunit $1 from Clostridium thermocellum YS.";

FEMS Microbiol. Lett. 78:181-186(1992).

-! FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE CELLULOSET OF THE CATALYTIC DOMAINS OF THE CELLULOSENT THROUGH THE BINDING OF THE NINE REPEATED DOMAINS WITH THE DUPLICATED SEGMENT OF 24 AMINO ACIDS PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSOME.

-!- SUBCELLULAR LOCATION: Cell surface.

-!- DOMAIN: THE COMPONENTS OF THE CELLULOSOME.

-!- DOMAIN: THE COMPONENTS OF THE CELLULOSOME.

-!- SIMILARITY: Contains at least 3 cohesin domains.
                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-0CT-1994 (Rel. 30, Created)
01-0CT-1994 (Rel. 30, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/SL)
(Cellulose integrating protein B) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clostridium thermocellum.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
                                                              1,
      Length 879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=93146373; PubMed=1490597;
Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
Gilbert H.J.;
                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Probom; PD001947; CBD 37 1.
PROSTIE; PS00018; EF HAND; UNKNOWN 1.
PROSTIE; PS00448; CLGS CELLULOSOME RPT; 2.
Cellulose degradation; Cell wall; Glycoprotein; Repeat.
NON TER 1 1 COHESIN 1.
DOMĀIN 81 93 LINKER (PRO/THR-RICH).
   DB 1;
35.2%; Score 165.5; DB 1
33.0%; Pred. No. 1.4e-10;
iive 26; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 772 AA.
                                                                                                                                                                                                                                                                                             88
                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR002102; Cohesin.
InterPro; IPR002105; Dockerin 1.
InterPro; IPR00246; EF-hand.
Pfam; PF00494; CBM 3; 1.
Pfam; PF00404; Cohesin; 3.
Pfam; PF00404; Dockerin 1; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; X68233; CAA48312.1; -. HSSP; Q06851; 1NBC.
   Query Match 35.23
Best Local Similarity 33.03
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Clostridium.
```

us-09-917-376-4.rsp

```
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JORN
JORN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TJRN
 9
                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fujino T., Beguin P., Aubert J.-P.;
"Organization of a Clostridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface.";
[3] Bacteriol. 175:1891-1899(1993).
                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                      "Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=97076134; PubMed=8918451;
TOTTO J., Lamed R., Chirino A.J., Morag E., Bayer E.A., Shoham Y.,
                                                                                                                                                                                                                                                                       01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Cellulosomal scaffolding protein A precursor (Cellulosomal glycoprotein S1/SL) (Cellulose integrating protein A) (Cohesin).
                                                                                                                                                                                                                                                                                                                                  Clostridium thermocellum.
Bacteria, Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
                                                                                                              29; Indels 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose."; EMBO J. 15:5739-5751(1996).
                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43.
STRAIN=ATCC 77405 / DSM 1237,
MEDLINE-93302508; PubMed=8316083,
Gerngross U.T., Romaniec M.P.M., Kobayashi T., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "A cohesin domain from Clostridium thermocellum: the crystal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            structure provides new insights into cellulosome assembly.";
Structure 5:381-390(1997).
                                                                                          Score 165; DB 1; Length 772; Pred. No. 1.4e-10;
                                                                      BBF06DE5E094FE10 CRC64;
          LINKER (PRO/THR-RICH).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361
MEDLINE=98022914; PubMed=9402065;
                   CELLULOSE-BINDING.
LINKER (PRO/THR-RICH)
                                                                                                                                                                                               336 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
MEDLINE=97238934; PubMed=9083107;
                                                                                                                                                                            61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522
                                                                                  35.1%; Scc...
36.1%; Pred. No. 1...
23; Mismatches
                                                                                                                                                                                                                                                  PRT; 1853 AA.
                                       COHESIN 3.
DOCKERIN 1.
DOCKERIN 2.
 COHESIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [2]
SEQUENCE OF 1820-1853 FROM N.A.
MEDLINE=93209931; PubMed=8458832;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mol. Microbiol. 8:325-334(1993).
                                                                                                                                                                                                                                                                      (Rel. 29, Created)
                                                                       82491 MW;
                                                                                                               Conservative
                                                                                                                                                                                                                                                  STANDARD;
240
272
439
461
607
733
                                                                        772 AA;
                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                     01-JUN-1994 (
01-NOV-1997 (
28-FEB-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Steitz T.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                             Demain A.L.;
                                                                                                                                                                                                                                                                                                                                                       Clostridium.
                                                                                                                 35;
                                                                                                                                                                                                                                      CIPA CLOTM
ID CIPA CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                           homology."
                                                                        SEQUENCE
                                                                                                                                                                                                                                                             006851;
DOMAIN
DOMAIN
DOMAIN
                              DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Frolow
                                                             DOMAIN
                                                                                                               Matches
                                                                                                                                                       유
                                                                                                                                                                                                                                                            ሯ
                                                                                                                                                                                               요
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                CELLULOLYTIC ENZYMES.
---- SUBCELLULAR LOCATION: Cell surface.
---- DOMAIN: THE COHESIN DOMAINS BIND TO THE DOCKERIN DOMAIN BORN BY
THE CATALYTIC COMPONENTS OF THE CELLULOSOME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DR PDB; INBC; 26-SEP-0.

DR InterPro; IPR001956; CBL DB 3.

DR InterPro; IPR001956; CelTul bind.

DR InterPro; IPR002102; Cohesin.

DR InterPro; IPR002104; Cohesin.

DR InterPro; IPR002104; Dockerin_1.

DR Pfam; PF00942; CBM_3; 1.

DR Pfam; PF00944; Dockerin.; 9.

DR Probom; P0001947; CBD_3; 1.

DR Probom; P0001947; CBD_3; 1.

DR PROSITE; PS00018; EF HAND; UNKNOWN 1.

NR PROSITE; PS00448; CLÖS_CELLULOSOME_RPT; 2.

W 3D-structure.

T SIGNAL

T SIGNAL

T SIGNAL
Tavares G.A., Beguin P., Alzari P.M.;
"The crystal structure of a type I cohesin domain at 1.7-A resolution.";
J. Mol. Biol. 273:701-713(1997).
-!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME.
PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF TI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULOSE-BINDING (BY SIMILARITY).
LINKER (PRO/THR-RICH).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULOSOMAL SCAFFOLDING PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LINKER (PRO/THR-RICH).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REF. 1).
                                                                                                                                                                                                                    -!- SIMILARITY: Contains 9 cohesin domains.
-!- SIMILARITY: Contains 2 dockerin domains.
                                                                                                                                                                                                                                                                                                                                                                                                                  or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COHESIN 3.
COHESIN 4.
COHESIN 5.
COHESIN 6.
COHESIN 7.
COHESIN 9.
COHESIN 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COHESIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOCKERIN 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; L08665; -; NOT ANNOTATED_CDS.
EMBL; X67506; CAA47840.1; -.
PIR; S36859; S36859.
PDB; 1ANU; 23-JUL-97.
PDB; 1AOH; 08-JUL-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28
1853
182
322
363
522
522
559
704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  031
196
361
361
526
690
690
814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  191
196
206
211
221
226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1548
1791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1824
1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  213
224
227
236
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                59
                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GUNV_ERWCA STANDARD; PRT; 505 AA.
047096;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V)
(Cellulase V).
                       01-NOV-1997 (Rel. 35, Created)
1-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last amotation update)
Endoglucanase N precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase N)
                                                                                              Erwinia carotovora.
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 444;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444 ENDOGLUCANASE N.
168 PROTON DONOR (BY SIMILARITY)
2556 NUCLEOPHILE (BY SIMILARITY).
48300 MW; FA7E4179004CBB43 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00150; cerrand.,
Probom; PD001947; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
31 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31.8%; Score 149.5; DB 1.38.6%; Pred. No. 3.8e-09; tive 19; Mismatches 30
  444 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DWANVGPNNIVTSTGTPAASTDKANRYV 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR001956; CBD 3.
Interpro; IPR001956; Cellul bind.
Interpro; IPR001547; Glyco_hydro_iPfam; PP00942; CBM 3; 1.
Pfam; PP00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; L39788; AAC37033.1; -. HSSP; O85465; 1A3H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444
168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            444 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
hes 34; Conserv
                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                        hydrolases).
                                                                                                                                   NCBI_TaxID=554;
                                                                          Cellulase N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GUNV ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
  유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8844488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35.1%; Score 165; DB 1; Length 1853; 36.1%; Pred. No. 3.9e-10; tive 23; Mismatches 29; Indels 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA; 196831 MW; 3ABDDC03ABFC5372 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          424 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                       1224
1229
1232
1242
1247
1249
1257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1292
1299
1302
                                                                                                                                                                                           431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
nes 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1853
                                                313
313
369
381
385
387
389
401
404
418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1263
1272
1279
1283
1289
1293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1301
1306
1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1313
1324
1329
1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S:
Matches 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
```

RESULT 12 GUNN_ERWCA

8 g ò 셤

```
hydrolases).
                                                                                 NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                        Erwinia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P50900;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 T 15
CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GUX2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
GUX2_CL
    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OSSIGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GUNW ERWCA STANDARD; PRT; 504 AA.
059335;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FRB-2003 (Rel. 41, Last annotation update)
Endoglucanase VI precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase VI)
(Cellulase VI).
                                                                                                                                                                                                                                                                                                                                                                            optimum about 42 degrees Celsius.
SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                                                                                                                                       Mol. Gen. Genet. 241:341-350(1993).
-!- FUNCTION: Endoglucanase with some exoglucanase activity.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- SUBCELLULAR LOCATION: Secreted.
-!- MISCELLANBOUS: Has a pH optimum of about 7:0 and a temperature
                                                                                                                                     STRAIN=SCRI193;
MEDLINE=94067016; PubMed=8246888;
Cooper V.J.C., Salmond G.P.C.;
"Molecular analysis of the major cellulase (CelV) of Erwinia carotovora: evidence for an evolutionary 'mix-and-match' of enzyme
                                        Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKER.
CELLUJOSE-BINDING (BY SIMILARITY).
PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31.8%; Score 149.5; DB 1; Length 505;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProDom; PD001947; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Pred. No. 4.4e-09; 19; Mismatches 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
ENDOGLUCANASE V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | | | | | | | | | : | | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CATALYTIC.
                                                            Enterobacteriaceae; Pectobacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP, 085465; IA3H.
InterPro, IPR001956; CBD 3.
InterPro, IPR008965; Cellul bind.
InterPro, IPR001547; Glyco_hydro_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54900 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00942; CBM 3; 1.
Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X76000; CAA53592.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 38.6 es 34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      505
334
352
505
168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR, S39962; S39962.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32
335
353
168
168
256
256
                        carotovora.
                                                                                                                     SEQUENCE FROM N.A.
                                                                               NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                   hydrolases)
                                                                                                                                                                                                                                      domains.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT SITE
ACT SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
GUNW_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              884444G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an enail to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                               Mae A., Heikinheimo R., Palva B.T.; "Structure and regulation of the Erwinia carotovora subspecies carotovora SCC3193 cellulase gene celV1 and the role of cellulase in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5; Gaps
                                                                                                                                                                                                                                                                                                                      phytopathogenicity.";

Mol. Gen. Genet. 247:17-26 (1995).

-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic

-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic

linkages in cellulose, lichenin and cereal beta-D-glucans.

-!- SUBCELLULAR LOCATION: Secreted.

-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Exoglucanase II precursor (EC 3.2.1.91) (Exocellobiohydrolase II)
(1,4-beta-cellobiohydrolase II) (Avicelase II).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULOSE-BINDING (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31.4%; Score 147.5; DB 1; Length 504; 37.5%; Pred. No. 7.4e-09; Live 19; Mismatches 31; Indels 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTON DONOR (BY SIMILARITY) NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP, 085465; 1A3H.
InterPro; IRR001956; CBD 3.
InterPro; IRR001956; Cellul bind.
InterPro; IRR001847; Glyco_hydro_5.
Pfam; PF00942; CBM 3; 1.
Pfam; PF00150; cellulase; 1.
Propom; PB00150; cellulase; 1.
PROSITE; RS00659; GLYCOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0D7ECF74781565FA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ENDOGLUCANASE V1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       411 DWANVGPNNIVTSTGTPAASTDKANRYV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LINKER
                                                                                                                                                                                                         MEDLINE=95231512; PubMed=7715600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54963 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X79241; CAA55823.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clostridium stercorarium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             504
334
352
504
168
256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; S54744; S54744.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      335
353
168
168
256
264 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
carotovora.
                                                                                                                                          SEQUENCE FROM N.A.
```

```
ON NEBL TEXTD-1510,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

RE STAIN-NOTED 11754,

R
```

Db 822 WASIGSSNVTGTFVKMDGATTGADYYLE 849
Search completed: May 11, 2004, 12:07:29
Job time: 4.42639 sees

8 8 8

61 WAAMGCGNIRASFGSVNPATPTADTYLO 88

```
PESTIT
                                                                                                       May 11, 2004, 11:45:55; Search time 10.385 Seconds (without alignments) 2704.020 Million cell updates/sec
                                                                                                                                                                                                   470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX 89
                                                                                                                                                                                                                                                                                                                                                 1017041
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                           1017041 segs, 315518202 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPTREMBL_25:*

1: sp_archea:*
2: sp_bacteria:*
4: sp_tungi:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vortebrate:*
14: sp_unclassified:*
15: sp_vortebrate:*
                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                  US-09-917-376-4
                                                                                                                                                                                Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                           Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description		Q82qf2 streptomyce	Q7x2n2 thermomonos	Q9rfx6 caldibacill	Q9rfx5 caldibacill	Q91818 caldibacill		Q9z4il bacillus sp		024820 thermophili			Q9x3p6 caldicellul		Q9aqh0 caldicellul	052374 caldicellul	Q93tj6 bacillus su
SUMMARIES	1		Q82QF2	Q7X2N2	Q9RFX6	Q9RFX5	Q9L8L8	Q8KKF7	Q9Z4I1	Q9AQG4	024820	Q9X3P5	Q9AQG7	99хзъ6	P96311	ОЭАОНО	052374	93176
	80	}	16	0	N	0	0	7	7	7	7	~	7	7	7	0	0	7
	% Query Match Length DB		741	919	170	930	921	1001	997	1751	1000	1770	261	1426	1711	966	1779	499
	% Query Match		54.7	48.6	48.0	46.1	45.6	44.6	43.5	41.4	41.0	41.0	40.7	40.7	40.6	40.5	40.5	40.4
	Score		257	228.5	225.5	216.5	214.5	209.5	204.5	194.5	192.5	192.5	191.5	191.5	191	190.5	190.5	190
	Result No.		н	7	m	4	S	9	7	89	σ	10	11	12	13	14	15	16

0931d0 bacillus su 052731 bacillus sp 045532 bacillus sp 045430 bacillus sp 08766 bacillus sp 087810 acetivibrio 097875 streptomyce 091318 clostridium 074386 bacillus 1i 09fdj9 bacteroides 09774 clostridium 045936 clostridium 045936 clostridium 062830 clostridium 062830 clostridium 062830 clostridium 062935 clostridium 06932 clostridium 06932 clostridium 06932 clostridium 0693119 bacillus sp 090642 ginglymosto	Q8jga9 ginglymosto Q9y3z2 homo sapien Q9y3z2 homo sapien Q913j2 clostridium O94858 homo sapien Q8izf2 homo sapien P91006 caenorhabdi Q8yru7 anabaena sp Q9gyg5 caenorhabdi Q99yg5 caenorhabdi
2 0931D0 2 052731 2 045332 2 045332 2 045332 2 045430 2 098PL0 16 097874 2 097836 2 097836 2 097836 2 097836 2 097836 2 097836 2 097836 2 045396 2 045396 2 045396 2 045396 3 046392 2 045396 3 046392 2 046392 3 046392 3 046392 3 046392 3 046392 3 046392 3 046392	13 Q8UGA9 4 Q9Y3Z2 2 Q9L3JZ 4 Q948S8 4 Q8IZF2 5 P91006 16 Q8YRU7 5 Q9GYGS 12 Q993M3
508 499 499 501 501 1945 6473 887 705 1168 1168 1168 1168 1168 1168 1168 116	121 618 660 986 1300 547 1428 511
4.00 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	16.12
1990 1870 1870 1870 1870 1870 1870 1870 187	76.55 75.55 75.55 75.55 73.55 73.55 73.55
11 11 11 11 11 11 11 11 11 11 11 11 11	. w w w 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

ALIGNMENTS

									stales;						ni C.,	sonoe T.,		comyces	ıry						() ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;	n H., Sniba T.,		the industrial				lycosyl; lEA.			
	741 AA.			Last sequence update)	Last annotation update)	sidase.			Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales	Streptomyces.			STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;		Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi	., Nakazawa H., Os	Kikuchi H., Shiba T., Sakaki Y., Hattori M.;	"Genome sequence of an industrial microorganism Streptomyces	avermitilis: deducing the ability of producing secondary		Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).			STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;		Shinose M., Kikuchi H., Shiba T.,		"Complete genome sequence and comparative analysis of the industrial	Γα			Finydrolase activity, hydrolyzing U-giycosyi	Ottem; ten:		
	PRT;			, Last seq	, Last ann	a-cellobio			Actinobact	cetaceae;			7 / NCIMB	1572948;	wa J., Han	Horikawa H	ki Y., Hat	ustrial mi	ability of		.A. 98:122			7 / NCIMB	2692562;		ra S.;	and compar	dvermitti	1(2003).		e activity	race meran	(50 3.	. DIIIG 1
	PRELIMINARY;		(TrEMBLrel. 24,	(TrEMBLrel.	(TrEMBLrel.	se 1,4-l	557.	Streptomyces avermitilis.	tinobacteria;	Streptomycineae; Streptomycetaceae; Streptomyces.	3903;	M.N.A.	80 / ATCC 3126	7403; PubMed=1	eda H., Ishika	Takahashi Y.,	Shiba T., Saka	ence of an ind	deducing the		Acad. Sci. U.S		M N.A.	80 / ATCC 3126	MEDLINE=22608306; PubMed=12692562;	Ikeda H., Ishikawa J., Hanamoto A.,	Sakaki Y., Hattori M., Omura S.;	nome sequence	microorganism streptomyces avermitties;	NAC. Blocechol. Z1:5Z6-531(Z0U3).	23; BACGG201:1	53; F:hydrolase activity, hydrolyz	GO; GO: UOUSS/S; F: Calbonyur	1PR001936; CBD 3	11100 COCOOCI
LT 1 F2	82QF2			01-JUN-2003	01-OCT-2003	Putative cel	GUXA1 OR SAV557.	Streptomyces	Bacteria; Ac	Streptomycin	NCBI_TaxID=33903;	SEQUENCE FROM N.A.	STRAIN=MA-46	MEDLINE=2147	Omura S., Ike	Shinose M.,	Kikuchi H.,	"Genome segue	avermitilis:	metabolites.";	Proc. Natl.	[3]	SEQUENCE FROM N.A.	STRAIN=MA-46	MEDLINE=2260	Ikeda H., IS	Sakakı Y., H	"Complete ger	microorganish	Mac. Blocech	EMBL; AFOUSO.	GO; GO: 0004553;	40; 40:00039 Tatorbro. 10	Interpro; IP	
RESULT ORZOFZ	£	AC	DŢ	F L	텀	DE	N	SO	ဗ	ဗ	X N	RP	RC C	ΚX	RA	RA	RA	RT	RI	RŢ	RL	RN N	RP	Z Z	X	Z:	Æ	# E	Z ;	¥ 5	Š i	E E	5 5	ž 2	Š

ខ

```
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Caldibacillus.
NCBI_TaxID=74586;
                                SEQUENCE FROM N.A. MEDLINE-20120520; PubMed=10653733; MEDLINE-20120520; PubMed=10653733; Sunna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.; Sunna A., Gibbs M.D., Chin Lidomain beta-1,4-mannanase from "A gene encoding a novel multidomain beta-1,4-mannanase from Caldibacillus cellulovorans and action of the recombinant enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=20120520; PubMed=10653733; Valson P.J., Bergquist P.L.; Subna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.; A gene encoding a novel multidomain beta-1,4-mannanase from Caldibacillus cellulovorans and action of the recombinant enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
48.0%; Score 225.5; DB 2; Length 170;
Best Local Similarity 47.7%; Pred. No. 8.2e-17;
Matches 41; Conservative 17; Mismatches 27; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                33 POTENTIAL.
930 MULTIDOMAIN BETA-1,4-MANNANASE.
101576 MW; 0086638D54D1A2CC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE 170 AA; 18493 MW; 7AC9D33F44E3A0B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  46.1%; Score 216.5; DB 2; 46.5%; Pred. No. 6.2e-15;
                                                                                                                                                                                                                                                                                         GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001956; CBD 3.
InterPro; IPR008965; Cellul_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Appl. Environ. Microbiol. 66:664-670(2000).
EMBL; AF163837; AAF22274.1; -.
HSSP; Q06851; INBC.
                                                                                                                                                                                                       Appl. Environ. Microbiol. 66:664-670(2000)
EMBL; AF163837; AAF22273.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Multidomain beta-1,4-mannanase precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProDom; PD001947; CBD<sup>-3</sup>; 2.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001956; CBD 3.
InterPro; IPR008965; CelTul bind.
InterPro; IPR004302; Chitin binding 3.
InterPro; IPR001647; Glyco hydro 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY_2000 (TrEMBLrel. 13, Created)
01-MAY_2000 (TrEMBLrel. 13, Last seq
01-OCT-2003 (TrEMBLrel. 25, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00942; CBM 3; 2.
Pfam; PF00150; cellulase; 1.
Pfam; PF03067; Chitin_bind_3; 1.
                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00942; CBM 3; 1.
ProDom; PD001947; CBD 3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                        Hypothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                             kraft pulp.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             kraft pulp.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 4
        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DDT TO DDT TO DD TO THE WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A WAY A 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Posta K., Beki E., Kukolya J., Hornok L.;
"Phylogenetic relationships of Tf cel5B, a new endoglucanase encoding gene from Thermobifida fusca.";
Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Hypothetical protein (Fragment).
Galdibacillus cellulovorans.
Bacteria, Firmicutes; Bacillales, Alicyclobacillaceae; Caldibacillus.
NCBI_TAXID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3;
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          48.6%; Score 228.5; DB 2; Length 616; 52.3%; Pred. No. 1.8e-16; Live 16; Mismatches 22; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                  Length 741;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptosporangineae; Nocardiopsaceae; Thermobifida.
                                                                                                                                                                                                                                                                                                                                                                                                                                      25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AY298814; AAP56348.1; -.
SEQUENCE 616 AA; 67701 MW; 24FFC1EA1A3F5639 CRC64;
                                                                                                                                                                                                                                                                                                                    741 AA; 77396 MW; 79404B40B2B4A7AF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                         54.7%; Score 257; DB 16; 56.3%; Pred. No. 1.7e-19;
InterPro; IPR008957; FN_III-like.
InterPro; IPR001524; Glyco_hydro_6.
Pfam; PF00942; CBM 3; 1.
Pfam; PF00041; Glyco_hydro_6; 1.
Pfam; PF001341; Glyco_hydro_6; 1.
PRINTS; PR00733; GLHYDRLASE6.
ProDom; PD001947; CBM 3; 1.
ProDom; PD003733; Glyco_hydro_6; 1.
SWART; SM00060; FN3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             616 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              170 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                    13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651 AAVGCSNVSLKVVPLTTPVPGADAYLE 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 AAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2003 (TrEMBLrel. 25, 01-OCT-2003 (TrEMBLrel. 25, 01-OCT-2003 (TrEMBLrel. 25,
                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 56.39
les 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Thermomonospora fusca.
                                                                                                                                                                                                                                                                                               Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. STRAIN=TM51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=2021;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Endoglucanase.
CELSB.
                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q7X2N2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              O9RFX6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
```

RESULT 2 Q7X2N2

g

8

ò

6

Length 930;

RESULT 3
O9RFX6
ID O9RF
AC 09RF
DT 01-M
DT 01-M
DT 01-M
CO DE HYPC
OS Calc
OX NCB1

g

ઠે

62

Gaps

1;

Matches

RESULT 5

셤 ò

```
340 TGTLEVQYRNGGSSASGNAITPQFNLKNTGTTAIDLSKVKVRYYYFTKD-SAADMSFWCDY 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=BP-23;
MEDLINE=21129642; PubMed=11234960;
MEDLINE=21129642; PubMed=11234960;
Pastor F.L.J., Pujol X., Blanco A., Vidal T., Torres A.L., Diaz P.;
"Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its performance in pulp
                                                                                                                                                                                                                                                  Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.

Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.

--- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

REMBL; AJ488933; CAD32445.1; ---
RGO; GO:0016162; F:cellulose 1.4-beca-cellobiosidase activity; IEA.

GO; GO:0016156; F:cellulose 1.4-beca-cellobiosidase activity; IEA.

RO; GO:0015975; P:cellulose 1.4-beca-cellobiosidase activity; IEA.

RO; GO:0015975; P:cellulose 1.4-beca-cellobiosidase activity; IEA.

RICEPPO; IPR001956; CBD 3.

RICEPPO; IPR001956; FN III-like.

RICEPPO; IPR001956; FN III-like.

RICEPPO; IPR001956; GJYco_hydro_48.

RICEPPO; IPR001956; GJYco_hydro_48.

RICEPPO; IPR0011; GJYco_hydro_48; 1.

RPTOM; PR0011; GJYco_hydro_48; 1.

RPTOM; PR0011; GJYco_hydro_48; 1.

RPTODOM; PD001197; CBD 3; 1.

RPTODOM; PD0011903; GJYco_hydro_48; 1.

RPTODOM; PD0011903; GJYco_hydro_48; 1.

RRANTY; SM00060; FN3; 2.

RRANTY; SM00060; FN3; 2.

RRANTY; SM00060; FN3; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                 STRAIN=BP-23;
Sanchez M.M., Pastor F.I.J., Diaz P.;
"Paenibacillus sp. BP-23 family 48 cellulase. Cloning and performance
on cellulosic substrates.";
                                            Paenibacillus sp. BP-23.
Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIĞNAL 1 35 POTENTIAL.
CHAIN 36 1091 CELLULOSE 1,4-BETA-CELLOBIOSIDASE.
SEQUENCE 1091 AA; 118001 MW; 21EACCEB2E704478 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1091;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appl. microbiol. Biotechnol. 55:61-68(2001).
-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
EMBL; AJ133614; CAB38941.1; -
Cellulose 1,4-beta-cellobiosidase precursor (EC 3.2.1.91)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacillus sp. BP-23.
Bacteria; Firmicutes; Bacillales; Bacillacéae; Bacillus.
NCBI_TaxID=89769;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
44.6%; Score 209.5; DB 2;
Best Local Similarity 46.0%; Pred. No. 4.4e-14;
Matches 40; Conservative 21; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP, P26221, 1TF4.
GO; GO:0008810, F:cellulase activity, IEA.
GO; GO:0005215; F:transporter activity; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             997 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-1999 (TrEMBLrel. 10, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1999 (TrEMBLrel. 10, Last s
01-OCT-2003 (TrEMBLrel. 25, Last a
Cellulase precursor (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                        NCBI_TaxID=198119;
                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  refining."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9Z4I1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09Z4I1
  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                         62
                                                62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            . . .; IEA.
                                                3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYMPTRDGGSSTLVYNCDWA
    1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Caldibacillus.
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE=20519260; PubMed=11065373;
Sunna A., Gibbs M.D., Bergquist P.L.;
Sunna I. thermostable multidomain 1,4-beta-xylanase from
'Caldibacillus cellulovorans' and effect of its xylan-binding domain
on enzyme activity.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  45.6%; Score 214.5; DB 2; Length 921; 46.5%; Pred. No. 1e-14; ative 16; Mismatches 29; Indels 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Microbiology 146:2947-2955 (2000).
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GI-HYDROLASES).
EMBL, AF200304; AAF61649.1; -.
HSSP, Q06651; 1NBC.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl
GO; GO:00005975; P:carbohydrate metabolism; IEA.
  28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 33 POTENTIAL,
34 921 BETA-1,4-XYLANASE XYNA.
921 AA; 102380 MW; C5DDDIA7F7567413 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                         01-077-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-0CT-2003 (TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001000; Glyco_hydro_10.
Pfam; PF00942; CBM 3; 2.
Pfam; PF02018; CBM 4 9; 1.
Pfam; PF00131; Glyco_hydro_10; 1.
PRINTS; PR00134; GLHYDRLASE10.
ProDom; PD001947; CBD_3; 2.
SMART; SM00633; Glyco_10; 1.
Glycosidase; Hydrolase; Signal; Xylan degradation.
    17; Mismatches
                                                                                                                                                                                                                                                                                                               921 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 1091 AA
                                                                                                                                                                      830 WVGCSNLRGSLVKLTTGRTGADYYLE 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                           63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR008965; Cellul bind.
InterPro; IPR008965; Cellul bind like.
InterPro; IPR001000; Glyco_hydro_10.
                                                                                                                                                                                                                                                                                                                                                                                                                                  Beta-1,4-xylanase XynA precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (TrEMBLrel. 22, C
(TrEMBLrel. 22, L
(TrEMBLrel. 25, I
  40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
les 40; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-OCT-2002 (
01-OCT-2002 (
01-OCT-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                               Q91818
Q91818;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q8KKF7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q8KKF7;
```

DDT TALE TO THE TOTAL TO THE TALE TO THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE THE TALE TH

Best Loc Matches

g ò 셤 RESULT 6

Q8KKF7

61

1; Gaps

```
SEQUENCE FROM N.A. STRAIN=NA10;
                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=67756;
                                                                                                                                                                                                                                                                                                                                                         Beta-glucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .09X3P5
Q9X3P5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
                                                                                                                                                                                                                                        RESULT 9
                                                                                                                                                                                                                                                      024820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9X3P5
                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                             2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA. GO; GO:0005215; F:transporter activity; IEA. GO; GO:0005975; P:carbohydrate metabolism; IEA. GO:0006810; P:transport; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL; AF078042; AAK06394.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ou-vun-2001 (TrEMBirel. 17, Last sequence update)
01-OCT-2003 (TrEMBirel. 25, Last annotation update)
CelE.
Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
NNRT TANTELLOSIRUPTOR.
                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                          Ϊ;
                                                                                                                                                                                                                                                                                                                                                      43.5%; Score 204.5; DB 2; Length 997;
43.7%; Pred. No. 1.4e-13;
tive 21; Mismatches 27; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1751 AA; 192176 MW; 60178CBF3C00BE95 CRC64;
                                                                                                                                                                                                                                                                                                     1 34 POTENTIAL.
997 AA; 106927 MW; F20CB119D6410629 CRC64;
GO:0005975; P:carbohydrate metabolism; IEA.
GO:0006810; P:transport; IEA.
                           InterPro; IPR001956; CBD 3.
InterPro; IPR008965; CBD 3.
InterPro; IPR008965; CBD 111] bind.
InterPro; IPR008961; FN III.
InterPro; IPR001701; Glyco hydro 9.
InterPro; IPR001701; Glyco hydro 9.
InterPro; IPR0008928; Glyco trans 6hp.
InterPro; IPR000866; Lipocin_cytFaBP.
Pfam; PF00041; fn3; 2.
Pfam; PF00041; fn3; 2.
Pfam; PF00041; fn3; 2.
Probom; PD001947; CBD 3; 1.
PROSITE; PS000592; GLYCOSYL HYDROL F9 1; 1.
PROSITE; PS000593; GLYCOSYL HYDROL F9 2; 1.
PROSITE; PS000591; LIPOCALIN; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IRRO01956; CBD 3.
InterPro; IRRO01956; CBD 3.
InterPro; IRRO019055; CBD 111 bind.
InterPro; IRRO01701; Glyco hydro 9.
InterPro; IRRO01701; Glyco hydro 9.
InterPro; IRRO00566; Lipocln cytFABP.
Pfam; PRO0142; CBM 3; 4.
Prom; PRO01447; CBD 3; 3.
PROSITE; PS00592; GLYCOSYL HYDROL F9 1; PROSITE; PS00592; GLYCOSYL HYDROL F9 2; 1.
PROSITE; PS000594; GLYCOSYL HYDROL F9 2; 1.
PROSITE; PS00134; LIPOCANL HYDROL F9 2; 1.
PROSITE; PS00134; LIPOCANL HYDROL F9 2; 1.
PROSITE; PS00134; LIPOCANL HYDROL F9 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | :| |:: | :|||| ||||:: | 905 AQVGSANVQGMFVAVNPAKGTADTYVE 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
                   P:transport; IEA.
                                                                                                                                                                                                                                                                             Glycosidase; Hydrolase; Signal SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9AQG4 PRELIMINARY;
Q9AQG4;
01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 43.7
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=80339
                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
```

```
63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GO; GO:0003779; F:actin binding; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Gaps
                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Miyake K., Machida Y., Hattori K., Iijima S.; "Characterization of a multi-domain cellulase from an extremely thermophilic anaerobe strain NAIO." Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
                                                       1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
41.0%; Score 192.5; DB 2; Length 1000;
Best Local Similarity 43.7%; Pred. No. 2.9e-12;
Matches 38; Conservative 17; Mismatches 31; Indels 1;
  Length 1751;
                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1000 AA; 113265 MW; B9F659A56A752C6B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
Query Match 41.4%; Score 194.5; DB 2; Best Local Similarity 44.7%; Pred. No. 3.4e-12; Matches 38; Conservative 14; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001589; Actbind carmin.
InterPro; IPR001586; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001965; CelTul bind.
InterPro; IPR00150; Glyco_hydro_10.
InterPro; IPR001547; Glyco_hydro_5.
IPR00150; CelTulase; 1.
IPR00150; CelTulase; 1.
IPR00151; Glyco_hydro_10; IPR00151; Glyco_hydro_10; IPR00151; Glyco_hydro_10; IPR000151; Glyco_hydro_10; IPR000151; SW00633; Glyco_l0; 1.
IPR05ITE; P8000194; GLYCOSYL_HYDROL_F10; IPR05ITE; P800559; GlyCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1770 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     431 AQIGASNVTFKFVKLSSSVSGADYYLE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                   :| |: :| :: || ||: 737 IGASNVTFNFVKLSSGVSGADYYLE 761
                                                                                                                                                                                                                             64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HYDROLASES)
EMBL, AB008029; BAA22939.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JAN-1998 (TrEMBLrel. 05, 01-JAN-1998 (TrEMBLrel. 05, 01-OCT-2003 (TrEMBLrel. 25,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 thermophilic anaerobe NA10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Glycosidase, Hydrolase.
```

```
177 IGASNVTFNFVKLSSGVSGADYYLE 201
                                                                                                                                                                                                                                                           64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AF078737; AAD30364.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HYDROLASES).
                                                                                                                               Hydrolase.
                                                                                                                                                          SEQUENCE
                                                                                                                                        NON TER
NON TER
                                                                                                                                                                                                                                                                                                                                         09X3P6
                                                                                                                                                                                                                                                                                                                                09X3P6
                                                                                                                                                                                                                                                                                                            RESULT 12
Q9X3P6
  셤
                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                  GO; GO:0030246; F:carbohydrate binding; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
GO; GO:0005975; F:carbohydrate metabolism; IEA.
InterPro; IPR00356; CBD 3.
InterPro; IPR005584; CBD 1V.
InterPro; IPR005084; CBM 6.
InterPro; IPR003056; CBM 6.
InterPro; IPR008056; Cellul bind.
InterPro; IPR008965; Cellul bind.
InterPro; IPR008965; Callul bind.
InterPro; IPR008065; Callul bind.
                                                                                                         01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Glycosyl hydrolase 6 (Fragment).
Glycosyl hydrolase 6 (Fragment).
Bacteria: Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Caldicellulosiruptor.
NCBI_TAXID=80339;
                                        XYNA.
Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                       41.0%; Score 192.5; DB 2; Length 1770; 43.5%; Pred. No. 5.7e-12; tive 16; Mismatches 31; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                      1770 AA; 193641 MW; 8BAF1937D4926C92 CRC64;
(TrEMBLrel. 12, Created)
(TrEMBLrel. 12, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           261 AA
                                                                                                                                                                                                                                                                                                                                                                              SMART; SMO0606; CBD_IV; 1.
SMART; SM00633; GlyCo_l0; 1.
SMOSITE; PSO0591; GLYCOSYL_HYDROL_F10; 1.
GlyCosidase; Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1163 IGASNVTFNFVKLSSGVSGADYYLE 1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [1] -
SEQUENCE FROM N.A.
STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                  Pfam; PF00942; CBM 3; 3.
Pfam; PF02018; CBM 4 9; 2.
Pfam; PF03422; CBM 6; 1.
Pfam; PF00311; GlyCo hydro 10; 1.
PRINTS; PR00134; GLHYDRLASE10.
                                                                                                                                                                                                 EMBL; AF078737; AAD30363.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 43.5
ses 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                  FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                       PD001947;
                                                                             NCBI_TaxID=80339;
01-NOV-1999 (
01-NOV-1999 (
01-OCT-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                       ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09A0G7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

```
Hash; quesal; lust.

Resign (1909) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2000) (2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;

"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";

Curr. Microbiol. 40:333-340(2000).

BMBL, AF078040; AAK06391.1; -.

HSSP; Q06851; INBC.

GO; GO:0016787; F:hydrolase activity; IEA.

GO; GO:005975; P:carbohydrate metabolism; IEA.

InterPro; IPR001956; CBD 3.

InterPro; IPR001956; Cellul_bind.

PF0Dom; PD001947; CBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Caldicellulosiruptor sp. Tok/B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 261;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 1
261 261
261 AA; 28759 MW; 4771744A26A6AE04 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 40.7%; Score 191.5; DB 2; Best Local Similarity 43.5%; Pred. No. 7.3e-13; Matches 37; Conservative 15; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 1426 AA
```

```
Local Similarity
tes 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                    Q06851; 1NBC
                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                          NCBI_TaxID=80339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            052374
RESULT 14
Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                        Bacteria; Firmicutes; Clostridia; Clostridiales; Anaerocellum group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zverlov V., Mahr S., Riedel K., Bronnenmeier K.;
Zverlov V., Mahr S., Riedel K., Bronnenmeier K.;
"Properties and gene structure of a bifunctional cellulolytic enz
(Celh) from the extreme thermophile Anaerocellum thermophilum wit
separate glycosyl hydrolase family 9 and 48 catalytic domains.";
Microbiology 144:457-465(1998)

Microbiology 144:457-465(1998)

PIRI, 731337; 731337

HSSP, P26221; ITF4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001956; CBD 3.
InterPro; IPR008965; Cellul bind.
InterPro; IPR000556; Glyco_hydro_48.
InterPro; IPR001701; Glyco_hydro_9.
InterPro; IPR008928; Glyco_hydro_9.
InterPro; IPR008928; Glyco_trans_6hp.
Pfam; PF00942; CBM_3; 3.
Pfam; PF00912; Glyco_hydro_48; 1.
Pfam; PF00914; Glyco_hydro_48; 1.
Pfam; PF009159; Glyco_hydro_98; 1.
                                                                                    ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7
                                                     ; Score 191.5; DB 2; Length 1426;
; Pred. No. 5.7e-12;
15; Mismatches 32; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40.6%; Score 191; DB 2; Length 1711;
43.8%; Pred. No. 8e-12;
ive 18; Mismatches 30; Indels
             Glycosidase, Hydrolase. – CB3FDB85D09A863 CRC64; SEQUENCE 1426 AA; 157544 MW; 29B3FDB85D09A863 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1711 AA; 189979 MW; E3E987CEB9CD0C21 CRC64;
                                                                                                                                                                                                                                                                                                   (TrEMBLrel. 03, Created)
(TrEMBLrel. 03, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                          PRT; 1711 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DWAQIGASNVTFKFVKLSSSVSGADYYLE 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProDom; PD001947; CBD 3; 2.
ProDom; PD011903; Glyco_hydro_48; 1.
PROSITE; PS00592; GLYCOSYL_HYDROL_F9_1; 1.
PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
 PS00659; GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                   472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                                                                                                        64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=Z-1320;
MEDLINE=98154434; PubMed=9493383;
                                                        40.7%;
                                                                                                                                                                                                                                                                                                                                                .,4-beta-glucanase (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43.8%;
                                                                                                                                                                                                                                                                                                                                                                           Anaerocellum thermophilum
                                                                                    37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  39; Conservative
                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                       Query Match
Best Local Similarity
Matches 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=31899;
                                                                                                                                                                                                                                                                                                     01-MAY-1997 (
01-MAY-1997 (
01-OCT-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
 PROSITE;
                                                                                                                                                                                                                                                                                         P96311:
                                                                                                                                                                                                                                                                          P96311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                                          ò
                                                                                                                                          셤
                                                                                                                                                                        ઠ
                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Miorobiol. 40:333-340(2000).
EMBL; AF078038; AAK06388.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Caldicellulosiruptor sp. Rt69B.1.
Bacteria, Firmicutes, Clostridia, Clostridiales, Syntrophomonadaceae,
Caldicellulosiruptor.
NCBI_TaxID=70295,
                                                                                                                                                                                                                                                                                                   Glycosyl hydrolase 5 (Fragment).
Caldicellulosiruptor sp. Tok7B.1.
Bacteria, Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40.5%; Score 190.5; DB 2; Length 996; 43.5%; Pred. No. 4.8e-12; tive 15; Mismatches 32; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=Rt69B.1;
Morris D.D., Gibbs M.D., Ford M., Thomas J., Bergquist P.L.;
"Family 10 and 11 xylanase genes from Caldicellulosiruptor sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    996 996
996 AA; 108275 MW; 3C72B6ED22F3C614 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
                                                                                                                        01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOSE; WOODS, THE CONTROL OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF STATE OF ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1779 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :| |: :| :: || |: 903 IGASNVTFNFVKLSSGVSGADYYLE 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 MGCGNIRASFGSVNPATPTADTYLQ 88
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Family 10 xylanase (EC 3.2.1.8)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AF036924; AAB95326.1;
PIR; T31085; T31085.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF02012; BNR; 9. Pfam; PF00942; CBM 3; 1. ProDom; PD001947; CBD 3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
```

```
GO; GO:0030246; F:carbohydrate binding; IEA.

GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . .; IEA.

R GO; GO:0005975; P:carbohydrate metabolism; IEA.

InterPro; IPR001564; CBD_1V.

InterPro; IPR001564; CBM_C.

InterPro; IPR001564; CBM_C.

InterPro; IPR001565; CBM_C.

InterPro; IPR001565; CBM_C.

InterPro; IPR0018955; CelTul_bind.

InterPro; IPR0018955; CelTul_bind.

InterPro; IPR0018955; CelTul_bind.

InterPro; IPR001895; Gal_bind_like.

InterPro; IPR001895; Gal_bind_like.

InterPro; IPR001895; CBM_3; 3.

InterPro; IPR001895; CBM_1; 1.

InterPro; IPR001895; CBM_1; 1.

InterPro; IPR001895; CBM_1; 1.

InterPro; IPR001895; CBM_1; 1.

InterPro; IPR001895; CBM_1; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; CBM_2; 1.

InterPro; IPR001895; IPR001895; CBM_2; 1.

InterPro; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR001895; IPR0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2; Length 1779;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
40.5%; Score 190.5; DB 2; Length
Best Local Similarity 43.5%; Pred. No. 9.6e-12;
Matches 37; Conservative 14; Mismatches 33; Indels
                                     DY X TO DY BY THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE STATE OF THE
```

8 8 8

64 MGCGNIRASFGSVNPATPTADTYLQ 88
: | | : : | | : | | 1172 IGASNVTFNFVKLTSGVSGADYYLE 1196

Search completed: May 11, 2004, 12:09:28 Job time : 12.385 secs

```
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELECOMMONICATION INFORMATION:
TELEPRAK: 215-540-9200
TELEPRAK: 215-540-5818
TELEX: «Unknown.»
INPORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 43, Appl
Sequence 2, Appli
Sequence 9, Appli
Sequence 9, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 12, Appli
Sequence 12, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                               May 11, 2004, 12:04:16 ; Search time 4.3675 Seconds (without alignments) 1052.023 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, A Sequence 7, A Sequence 7, A Sequence 7, A Sequence 22, Sequence 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 47,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 10,
Sequence 12,
Sequence 7, 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 29,
Sequence 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                             470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-136-574A-47
US-09-136-574A-43
US-09-136-574A-43
US-07-186-58BB-2
US-09-033-537A-1
US-09-033-537A-1
US-09-033-159B-4
US-09-198-956-10
US-09-198-956-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-09-670-141-10
US-08-646-53-12
US-08-460-458-2
US-08-460-458-2
US-08-30-394A-2
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-006-632-7
US-09-009-161B-22
US-09-301-394A-22
US-09-301-394A-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389414 segs, 51625971 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                        OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Issued Patents AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match Length DB
                                                                                                                                                                                                                                                                                   US-09-917-376-4
                                                                                                                                                                                                                                                                                Title:
Perfect score:
Seguence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            195.5
194.5
191.5
191.5
165.5
165
165
165
165
165
165
165
165
165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
                                                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
```

Sequence 5, Appli Sequence 30, Appli Sequence 30, Appli Sequence 3, Appli Sequence 18, Appli Sequence 43, Appli Sequence 4640, Appli Sequence 38, Appli Sequence 31, Appli Sequence 32, Appli Sequence 32, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 2, Appli Sequence 26, Appli Sequence 26, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli Sequence 68, Appli	Wethods for see Containing Fabrics Using Truncated Compositions Center, P.O. Box 457	/CIP
US-09-118-319-5 US-08-315-400-36 US-09-301-593-30 US-09-341-587-3 US-09-341-587-3 US-09-310-593-18 US-09-310-593-18 US-09-328-352-464 US-09-328-352-464 US-09-328-352-464 US-09-423-439-38 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-09-687-538B-8 US-08-470-350B-2 US-08-470-350B-2 US-08-470-350B-2	ALIGNMENT 136574A Graham K. ige and eter W. positions and ating Cellulc lulase Enzyme and Howson se Corporate tte atible DOS DOS E. US/09/136,5	UMBER: US 08/932,571 September 19, 1997 CORMATION: LY E. NUMBER: 31,215 NUMBER: 1997US001/CIP INFORMATION:
13.6 428 13.6 472 4 13.5 13.5 13.4 472 4 13.4 472 4 13.4 472 4 453 4 13.3 13.3 485 1 13.3 13.3 485 1 13.3 13.3 485 1 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3 13.3	ication US 6 6 Farringto Anderson, Beighs, M Beighs, M Beighs, M Beighs, M Beighs, M WENTION: NVENTION: SEQUENCES: SEQUENCES: SPA SPA SPA SPA SPA SPA SPA SP	APPLICATION NUMBER: US 08 FILING DATE: September 19 ATTORNEY AGENT INFORMATION: NAME: Bak, Mary E. REGISTRATION NUMBER: 31,2 REFERENCE/DOCKET NUMBER:
228 239 330 331 331 332 333 334 335 444 443 443 444 443 444 444	RESULT 1 US-09-136-574A-47 Sequence 47, Appl Patent No. 629436 GENERAL INFORMA APPLICANT: TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I TITLE OF I	ATTORN N N N N N N N N N N N N N N N N N N

```
678 GVKVLYKNNETSASTGSIRPWFKIVNGGSSSVDLSRVKIRYWYTVDGDKPQSAV-CDWAQ 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 43:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 43.5%
Warches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-136-574A-43
                                                                                                                                                               US-09-136-574A-43
                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Treating Cellulose Containing Fabrics Using Truncated
                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                    Gaps
                                                                                   .,
H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 41.4%; Score 194.5; DB 3; Length 1751; Best Local Similarity 44.7%; Pred. No. 1.1e-12; Matches 38; Conservative 14; Mismatches 32; Indels 1;
                                         Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
                                                                                34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                          Anderson, Paige
Glabbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
                                       41.6%; Score 195.5; DB 3;
43.2%; Pred. No. 2.4e-13;
tive 15; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION «Unknown»
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                             61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                       60 WAQIGASNVTFNFVKLSSGVSGADYYLE 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                                                                                                                                                         Sequence 44, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 44:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Bak, Mary E.
                                                             Best Local Similarity 43.29
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 19477
                                                                                                                                                                                                                                                                                                                  RESULT 2
US-09-136-574A-44
US-09-136-574A-47
                                                                                                                                                                 d
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                임
```

```
Fabrics Using Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 GVKVQYKDNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                              Treating Cellulose Containing
Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: 1997US001/CIP TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40.7%; Score 191.5; DB : 43.5%; Pred. No. 1.8e-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US 08/932,571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY: linear MOLECULE TYPE: No. 6294366e SEQUENCE DESCRIPTION: SEQ ID NO: 43:
                                                                                                                                                                                                                    Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REGISTRATION NUMBER: 31,215
                                                                                                                                     Sequence 43, Application US/09136574A
                                                                                                                                                        Patent No. 6294366
GENERAL INFORMATION:
APPLICANT: Farrington, Graham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 1426 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEPHONE: 215-540-9200
```

RESULT 4

4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA

```
405 Lexington Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX: 212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TOPOLOGY: linear
US-09-033-537A-1
                      New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: D
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                        APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59580830 No. 5958083disk of No. 5958083th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36.6%; Score 172; DB 2; Length 700; ilarity 44.9%; Pred. No. 1.1e-10; Conservative 18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM FO Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Onishi, Masahiro
APPLICANT: Fich, Merete
APPLICANT: Toft, Annette Hanne
APPLICANT: Toft, Annette Hanne
APPLICANT: Sh lein, Martin
TITLE OF INVENTION: Prevention Of Back-Staining
TITLE OF INVENTION: In Stone Washing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             / Lambiris, Elias J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
PRIOR DATE: 19-JAN-1990
PRIOR DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, Elia
REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 3425.204-US
TELEFONNICATION INFORMATION:
TELEFONNICATION:
TELEFONNICATION:
TELEFONNICATION:
TELEFONNICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
               Sequence 2, Application US/07862588B Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1, Application US/09033537A Patent No. 5958083 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 700 amino acids TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New York
: U.S.A.
                                                                                                                                                                                                                                                                 New York
                                                                                                                                                                                                                                                                                                                                 10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-033-537A-1
                                                                                                                                                                                                                                                                                                            COUNTRY:
    US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
35.2%; Score 165.5; DB 2; Length 551;
Best Local Similarity 37.9%; Pred. No. 4.1e-10;
Matches 33; Conservative 22; Mismatches 29; Indels 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Yeda Research and Development Co. Ltd.
APPLICANT: Ramot University Authority for Applied
APPLICANT: Research and Industrial Development Ltd.
APPLICANT: Research and Development Ltd.
APPLICANT: Bayer, Edward A.
APPLICANT: Worag, Ely
APPLICANT: Wichek, Meir
APPLICANT: Wichek, Meir
APPLICANT: Lamed, Raphael
APPLICANT: Lamed, Raphael
APPLICANT: MODIFIED CELLULOSE-BINDING DOMAIN (CBD)
TITLE OF INVENTION: PROTEINS AND USE THEREOF
                                                                                                                           COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
OPERATING SYSTEM: DOS
OPTWARE: FASTERO FOR WINDOWS VERSION 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,537A
FILING DATE: 0.2-MAR-1998
CLASSIFICATION: 0.08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
ADTREST 419 Seventh Street N.W., Ste. 300
CITY: Washington
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                461 AQIGASNVSAAF--ANFTGSNTDTYVE 485
                                                                                                                                                                                                                                                                                                       PRICE APPLICATION DATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
APPLICATION NUMBER: PCT/DK96/00364
ATTORNEY/AGENT INPORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 AAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4492.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PCT-US95-13813-9
; Sequence 9, Application PC/TUS9513813
; GENERAL INFORMATION:
STATE: N. COUNTRY: U.S.A. 2IP: 10174
COMPUTER READABLE FORM: MEDIUM TYPE: Diskette COMPUTER: IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: 44
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ouery Match 35.1%; Score 165; DB 5; Length 167; Best Local Similarity 36.1%; Pred. No. 1.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 165; DB 4; Length 476;
Pred. No. 3.9e-10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||: :| ||||:
373 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 409
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS.
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAM ------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Kauppinen, Markus
APPLICANT: Schulein, Martin
APPLICANT: Schnorr, Kirk
APPLICANT: Andersen, Lene
APPLICANT: Bjornvad, Mads
TITLE OF INVENTION: No. 6566114el Mannanases
FILE REFERENCE: 5440.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/339,159B
CURRENT FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PATENTIN VERSION 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Application US/09339159B Patent No. 6566114
GENERAL INFORMATION:
                                                                                                                                                                     ATTORNEY AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTATION UNDRER: 25,618
REFERENCE/DOCKET NUMBER: BAYER
TELECOMMUNICATION INFORMATION:
TELEFRAN: (202) 737-3528
TELEFAX: (202) 737-3528
TELEFAX: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 35.1%;
Best Local Similarity 36.1%;
Matches 35; Conservative 2:
                                                                                                                                                                                                                                                                                                                                                                                                                    STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT ORGANISM: Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 7
US-09-339-159B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-339-159B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

RESULT 8

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10; Gaps
                                                                                                                                   APPLICANT: SCHLILLEIN, Mals Erik K.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Ginnor, Kirk K.
APPLICANT: Schnorr, Kirk Pectin Degrading Enzymes From Bacillus
TITLE OF INVENTION: Licheniformis
TITLE OF INVENTION: Licheniformis
TITLE OF INVENTION: Licheniformis
TITLE APPLICATION NUMBER: US/09/198,956
CURRENT FILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: 1344/97
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-12-02
SARLIER FILING DATE: 1997-12-02
SOFTWARRE: FastSEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH: 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 35.1%; Score 165; DB 3; Length 493; Best Local Similarity 36.1%; Pred. No. 4.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Lange, Niels E.
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soran
APPLICANT: Glad, Sanne O. S.
APPLICANT: Rauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILE REFERENCE: 5378.200-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/198,955A CURRENT FILING DATE: 1998-11-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRIOR APPLICATION NUMBER: 1343/97
PRIOR PELICATION NUMBER: 1344/97
PRIOR PELICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1998-12-02
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-11-02
PRIOR FILING DATE: 1998-11-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 12, Application US/09198955A Patent No. 6187580
Sequence 10, Application US/09198956
Patent No. 6165769
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Andersen, Lene N.
Schulein, Martin
                                                                                       Lene N.
Martin
                                                                                    APPLICANT: Andersen, APPLICANT: Schulein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-198-956-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Fa
SEQ ID NO 12
LENGTH: 493
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                      331 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 389
                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                 Query.Match 35.1%; Score 165; DB 3; Length 493;
Best Local Similarity 36.1%; Pred. No. 4.1e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 35.1%; Score 165; DB 4; Length 493; Best Local Similarity 36.1%; Pred. No. 4.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels 10;
                                                                                                                                                                                                                                                                                                 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                     61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: NO. 6368843el Pectate Lyases
FILE REFERENCE: 5378 200-US
CURRENT APPLICATION NUMBER: US/09/694,531
CURRENT FILING DATE: 2000-10-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION NUMBER: 1343/97
PRIOR PELING DATE: 1997-11-24
PRIOR PLILING DATE: 1997-11-24
PRIOR PLILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR PELING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-11-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-694-531-12
                  ; ORGANISM: Clostridium thermocellum US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION NUMBER: 09/198,955
PRIOR FILING DATE: 1998-11-24
                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 12, Application US/09694531
Patent No. 6368843
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Andersen, Lene N. APPLICANT: Schulein, Martin. APPLICANT: Lange, Niels E. APPLICANT: Bjornvad, Mads E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Moller, Soren
Glad, Sanne O. S
                                                                                                                                                                                                                                                                                                                                                                                                          US-09-694-531-12
; TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                 ð
```

RESULT 11 US-09-670-141-10 ; Sequence 10, Application US/09670141

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10; Gaps
                                         APPLICANT: AGELSEN, Jens N. APPLICANT: ACCESS.
APPLICANT: Schulein, Martin APPLICANT: Bornvad, Mach E. APPLICANT: Bornvad, Mads B. APPLICANT: Bornvad, Mads B. TITLE OF INVENTION: Dectin Degrading Enzymes From Bacillus TITLE OF INVENTION: Licheniformis FILE REFERENCE: 5372.200-US CURRENT APPLICATION NUMBER: US/09/670,141
CURRENT APPLICATION NUMBER: 13040-09-26
PRIOR PILING DATE: 1990-11-24
PRIOR FILING DATE: 1990-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH: 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 35.1%; Score 165; DB 4; Length 493; Best Local Similarity 36.1%; Pred. No. 4.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: GOLLEY, SOLEN
APPLICANT: GOLLEY, SOLEN
APPLICANT: GLAG, Sanne O. S.
APPLICANT: GLAG, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kongsbak, Lare
TILLE OF INVENTION: No. 6677147el Pectate Lyases
TILLE REFERENCE: 5378 200-02
CURRENT APPLICATION NUMBER: US/09/198,955
PRIOR APPLICATION NUMBER: US/09/198,955
PRIOR FILING DATE: 1998-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR PUBLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR PUBLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR PLING DATE: 1998-11-02
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR PLING DATE: 1998-11-02
PRIOR APPLICATION NUMBER: 09/184,217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 12, Application US/10072152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: Bacillus licheniformis US-09-670-141-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mads E
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lange, Niels E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Moller, Soren
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                jornvad,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Patent No. 6677147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -10-072-152-12
```

ò

8 셤

```
COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                        62 AAMGCGNIRASFGSVNPATPTADTY
                                                                                                                                                                     Sequence 2, Application US/08048164A Patent No. 5496934 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2, Application US/08460462
Patent No. 5670623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, No. METHODS OF
NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                         PENNIE & EDMONDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : 162 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 29.9*
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          inear
                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: New York
STATE: New York
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: Shosey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: U
ZIP: 10036
                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-048-164A-2
                                                                                                                                                      US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
                                                                                                                           RESULT 14
                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. 405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                          10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2; Length 531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
CORRESPONDENCE: 7
CORRESPONDENCE ADDRESS:
                                                                                                                             35.1%; Score 165; DB 4; Length 493; 36.1%; Pred. No. 4.1e-10; tive 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Floppy disk

COMPUTER: Ploppy disk

COMPUTER: Ploppy disk

COMPUTER: Ploppy disk

COMPUTER: Ploppy disk

COMPUTER: Ploppy disk

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/862,588B

FILING DATE: 19920727

CLASSIFICATION DATA:

APPLICATION NUMBER: DK 164/90

FILING DATE: 19-JAN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DK91/00013

FILING DATE: 18-JAN-1991

ATTONERY/AGENT INFORMATION:

NAME: Zelson, Steve T. / Lambiris, Elias J.

REGISTRATION NUMBER: 30,335 / 33,728

REGISTRATION NUMBER: 30,335 / 33,728
                                                                                                                                                                                                                                                                                                                                390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                     61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25.6%; Score 120.5; DB 2; 32.9%; Pred. No. 3.5e-05; tive 18; Mismatches 36;
; SEQ ID NO 12
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-10-072-152-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 7, Application US/07862588B Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELECOMMUNICATION INFORMATION: TELEPHONE: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: 212 867 0298 INFORMATION FOR SEQ ID NO: 7:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: AMINO ACID
                                                                                                                                                                          35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Marches 28; Conserv?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New York
: New York
: U.S.A.
                                                                                                                                                      Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                  Query Match
                                                                                                                                                                          Matches
```

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 ISSMSVEFYNSNKSAQINSITPIIKIINTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24.3%; Score 114; DB 1; Length 162; 29.9%; Pred. No. 4.2e-05; tive 22; Mismatches 34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARES PATENTIA Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,164A
FILING DATE: U4-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MAISTOCK, S. Leslie
REGISTRATION, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-003
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFRAX: (212) 790-9090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 AAMGCGN-----IRASF--GSVNPATPTADIYLQ 88
                                                                                                                                                                                                                                                                                                                  APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Soldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                            491 AQIGRTNVLLAF -- ANFTGSNTDTY 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET: 1155 Avenue of the Americas
86
```

```
CORRESPONDENCE ADDRESS:
ADDRESSES:
PRINTER: PRINTE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
COUNTRY: U.S.A.
COMPUTER: New York
COMPUTER: RADABLE FORM:
COMPUTER: RADABLE FORM:
COMPUTER: Batenil Ne Compatible
COMPUTER: IBM PC Compatible
COMPUTER: TOWN UNDRY:
COMPUTER: Batenil Ne Compatible
COMPUTER: Date Not Compatible
COMPUTER: STATE: New York
COMPUTER: STATE: New York
COMPUTER: IBM PC Compatible
COMPUTER: SAFETING NATA:
COMPUTER: SAFETING NATA:
COMPUTER: SAFETING NATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: Labelle
REGISTRATION NUMBER: 18,002
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING NATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
FILING DATE: Labelle
REGISTRATION NUMBER: 18,003
```

3 TSSMSVEFYNSNKSAQTNSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH 61

셤

Search completed: May 11, 2004, 12:11:21 Job time : 5.3675 secs

```
1346
1346
1346
1346
1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-5
190
180
165
165
165
165
155.5
                                                                                                                                      FEATURE:
   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                   May 11, 2004, 12:09:36 ; Search time 10.385 Seconds (without alignments) 2378.773 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4
Sequence 4
Sequence 1
Sequence 5
Sequence 5
Sequence 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence (
Sequence 1
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Cgn2_6/ptodata/2/pubpaa/USO7_PUBCOMB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO7_NEW_PUB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO6_PUBCOMB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO7_NEW_PUB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO8_NEW_PUB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO8_NEW_PUB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
| Cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
               GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-5
US-10-155-400-5
US-09-917-376-4
US-09-917-378-4
US-09-917-378-1
US-09-917-384-5
US-09-917-384-6
US-09-917-384-6
US-09-917-384-6
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-10-155-400-1
US-10-155-400-1
                                                                                                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                              1140673 segs, 277566755 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Published Applications AA:*
                                                                                                                                                                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                          protein search, using sw model
                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                         seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match Length DB
                                                                                                                                                                US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150
1043
1043
1228
1228
                                                                                                                                                                   Title:
Perfect score:
                                                                                                                                                                                                                                  Scoring table:
                                                                            OM protein
                                                                                                                                                                                                                                                                                                                                         Minimum DB
Maximum DB
                                                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                          Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š.
```

```
Sequence 51483, A Sequence 6, Appli Sequence 22, Appli Sequence 2, Appli Sequence 49900, A
                                                                                                                                                                                                                                            Sequence 223, App
Sequence 53, Appl
Sequence 53, Appl
Sequence 101, App
                                                                                                                                                                                                                                                                                                                                    Sequence 3, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 20, Appl
Sequence 20, Appl
Sequence 61, Appl
Sequence 61, Appl
Sequence 61, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 15685,
Sequence 15685,
                                                                                                                 Appli
Appli
                                                                                                                                                                                                                                                                                                                        Sequence 406, App
                                      Sequence 1
Sequence 1
Sequence 1
Sequence 2
                                                                                                               Sequence 2
Sequence 4
Sequence 5
5 US-10-369-493-23151

4 US-10-185-990-10

2 US-10-655-433-12

3 US-10-655-433-12

3 US-10-072-152-12

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-4

0 US-09-784-554B-3

0 US-00-782-122A-49900

0 US-09-80B-898-22

0 US-10-281-122A-49900

0 US-09-10-312-312-53

0 US-10-312-312-53

0 US-10-312-312-54

0 US-10-312-312-54

0 US-09-747-835A-24

0 US-09-747-835A-61

0 US-09-747-835A-61

0 US-09-747-835A-61

0 US-09-747-835A-61

0 US-10-312-312-61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-747-835A-25
US-10-312-312-25
US-10-424-599-196885
US-09-764-860-377
```

ALIGNMENTS

```
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                           APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HOMEN, TODD B.
APPLICANT: HIMMEL, TODD B.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REPERENCE: 40197.4US01
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 5
LENGTH: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 99.8%; Score 469; DB 12;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; OTHER INFORMATION: Carbohydrate binding domain US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Acidothermus cellulolyticus
Sequence 5, Application US/09917376
Publication No. US20040038334A1
GENERAL INFORMATION:
```

ö

```
61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER INFORMATION: Any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 100.0
Matches 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 100.
Matches 88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                슝
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.4US01
CURRENT FAPILICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATCHIN Ver. 2.1
SEQ ID NO 4
LENGTH: 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
0
                                                                                       APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, MILLIAM S.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
FITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: NREL 01-36A
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATCHIN VET. 2.1
SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99.8%; Score 469; DB 14;
100.0%; Pred. No. 1.5e-48;
live 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ch 99.8%; Score 469; DB 12; Similarity 100.0%; Pred. No. 1.5e-48; B8; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE:
, OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: Any amino acid OTHER INFORMATION: Carbohydrate binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Acidothermus cellulolyticus
                        Sequence 5, Application US/10155400 Publication No. US20030108988A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 4, Application US/09917376
Publication No. US20040038334A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 88; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 88; Conserv
                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRXWFIRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           , APPLICANT: ADNEY, WILLIAM S.

JAPPLICANT: DECKER, STEPHEN S.

APPLICANT: DECKER, STEPHEN R.

APPLICANT: HIMMEL, MICHAEL R.

TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 40197.7USO1

CURRENT APPLICATION NUMBER: US/09/917,378

CURRENT FILING DATE: 2001-07-28

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PATENTIN Ver. 2.1

SEQ ID NO 4

LENGTH: 154

TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANY, TOOD B.
APPLICANT: HIMMEL, TOOD B.
TITLE OF INVENTION: TELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TILE REFERENCE: NREL 01-36A
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 4
LENGTH: 89
TYPE: no.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate; OTHER INFORMATION: binding domain
US-09-917-378-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.8%; Score 469; DB 10; Length 154; 100.0%; Pred. No. 2.9e-48; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99.8%; Score 469; DB 14;
100.0%; Pred. No. 1.5e-48;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ) OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
```

ö

.. 0

5 C

ò

q

ઠ

```
1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRXWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: DECKER, STEPHEN R.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT FILING DATE: 2001-07-28
CURRENT FILING DATE: 201-07-28
SOFTWARE: PATENTING DATE: 201-07-28
SOFTWARE: PATENTING DATE: CENTRAL SOFTWARE: PATENTING DATE: CENTRAL SOFTWARE: PATENTING DATE: CENTRAL SOFTWARE: PATENTING DATE: CENTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: PATENTING DATE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWARE: CONTRAL SOFTWAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CELLULASE FROM ACIDOTHERMUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OTHER INFORMATION: Description of Artificial Sequence: Segment; OTHER INFORMATION: GuxA
US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                          Indels
Best Local Similarity 98.9%; Pred. No. 6.5e-48; Matches 87; Conservative 1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 466; DB 10;
Pred. No. 6.5e-48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88
                                                                                                                                                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                  61 WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5, Application US/09917383
Publication No. US2030104522A1
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 6, Application US/09917384 Publication No. US20030096342A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 98.9
Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-384-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         원
                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             455 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                   9
                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGSSSTLVTNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1, Application US/09917378
; Sequence 1, Application US/09917378
; Publication No. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: DING, SHI-YOU
; APPLICANT: DECKER, STEPHER R.
; APPLICANT: HIMMEL, MICHAEL E.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REPERENCE: 40197.7703
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 5, Application US/09917384

Publication No. US20030096342A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: ADNEY, WILLIAM S.

APPLICANT: HIWMEL, MICHAEL B.

APPLICANT: HIWMEL, MICHAEL B.

TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 40170.64081

CURRENT APPLICATION NUMBER: US/09/917,384

CURRENT FILING DATE: 2001-07-28

NUMBER OF SEC 1D NOS: 14

SOFTWARE: PATENTIN Ver. 2.1

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO S: 14

SEQ 1D NO 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99.8%; Score 469; DB 10; Length 762; 100.0%; Pred. No. 2e-47; tive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99.1%; Score 466;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-917-384-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
```

g ઠ 셤

9

Gaps

ö

```
LENGTH: 1228
                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
US-09-917-383-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-383-1
                                                                                                                                  US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-1
                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 원
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                   ö
                                                                                                                                                                         477 VSGGLKVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 477 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09917384

Publication No. US20030096342A1

GENERAL INPORMATION:

APPLICANT: DING, SH-YOU

APPLICANT: ADNEY, WILLIAM S.

APPLICANT: HIMMEL, MICHAEL B.

APPLICANT: HIMMEL, MICHAEL B.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 40170.64801

CURRENT APPLICATION NUMBER: US/09/917,384

CURRENT FILING DATE: 2001-07-28

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                     Sequence 6, Application US/09917383
Sequence 6, Application No. US20030104522A1
GENERAL INCPRATION:
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: UNMBER: US/09/917,383
CURRENT APPLICATION NUMBER: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
; OTHER INFORMATION: Description of Artificial Sequence: Segment of; OTHER INFORMATION: GuxA
US-09-917-384-6
                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURE:

OTHER INFORMATION: GuxA
US-09-917-383-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1043;
                                                                              Length 1043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                 0; Indels
                                                                          Score 466; DB 10;
Pred. No. 6.8e-47;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 / Match 99.1%; Score 466; DB 10; Local Similarity 98.9%; Pred. No. 6.8e-47; nes 87; Conservative 1; Mismatches 0;
                                                                                                                                                                                                                                                   WAAMGCGNIRASFGSVNPATPTADTYLO 564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88
                                                                                                                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                            99.1%;
98.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Artificial Sequence
                                                                    Query Match
Best Local Similarity 98.99
Matches 87, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGIH: 1043
                                                                                                                                                                                                                                                                                                                                           US-09-917-383-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                    ò
                                                                                                                                                                                       d
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                       584 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT; TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERNAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: UNMERR: US/09/917,383
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PALENTIN Ver. 2.1
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/09917376
Publication No. US20040038334A1
GENERAL INFORMATION:
APPLICANT: DING, SHILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: VINZANT, TODD B.
APPLICANT: VINZANT, TODD B.
APPLICANT: TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS; TITLE OF INVENTION: CELLUCAYTICUS
FILE REFERENCE: 40197.4 UŞ01
                                                                                        FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Segment of
OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: Description of Artificial Sequence: Segment OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99.1%; Score 466; DB 10; Length 1228; 98.9%; Pred. No. 8.3e-47; cive 1; Mismatches 0; Indels C
                                                                                                                                                                                                               Length 1228;
                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                               Score 466; DB 10;
Pred. No. 8.3e-47;
                                                                                                                                                                                                                                                           1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/09917383
Publication No. US20030104522A1
GENERAL INFORMATION:
                                                                                                                                                                                                                 99.1%;
98.9%;
SEQ ID NO 1
LENCTH: 1228
TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 98.9'
Matches 87; Conservative
                                                                                                                                                                                                                                     Best Local Similarity 98.9
Matches 87; Conservative
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                   869 VSGGVKVQYKANDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYMFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1. Application US/10155400

Publication No. US20030108988A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: VINZANT, TODD B.

APPLICANT: VINZANT, TODD B.

APPLICANT: VINZANT, TODD B.

TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS

TITLE OF INVENTION: TELLULOLYTICUS

FILE REPRESENCE: NREL 01-36A

CURRENT APPLICATION NUMBER: US/10/155,400

CURRENT APPLICATION NUMBER: US/10/155,400

CURRENT APPLICATION NUMBER: US/10/155,400

SOFTWARE: PATENTING DATE: 2002-10-22

NUMBER OF SEQ ID NOS: 7

SEQ ID NO 1

LENGTH: 957
                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.3%; Score 462; DB 14; Length 957; 97.7%; Pred. No. 1.9e-46; tive 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                98.3%; Score 462; DB 12; Length 957; 97.7%; Pred. No. 1.9e-46; Live 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   929 WAAIGCGNIRASFGSVNPAIPTADIYLO 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          929 WAAIGCGNIRASFGSVNPAIPTADIYLO 956
                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WAAMGCGNIRASFGSVNPATPTADTYLO 88
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SSCTWARE: Patentin Ver. 2.1
SSC ID NO 1
LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: Acidothermus cellulolyticus
                                                                                                                       TYPE: PRT
ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , Sequence 8100, Application US/10156761; Publication No. US20030119018A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCATION: (957)
CTHER INFORMATION: Any amino acid US-10-155-400-1
                                                                                                                                                                                                     LOCATION: (957)
OTHER INFORMATION: Any amino acid
                                                                                                                                                                                                                                                                                                                        86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
, APPLICANT: OMURA, SATOSHI
, APPLICANT: IKEDA, HARUO
                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 86; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 86; Conserv
                                                                                                                                                            FEATURE:
NAME/KEY: MOD RES
LOCATION: (957)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY: MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-10-156-761-8100
                                                                                                                                                                                                                                             US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 54.7%; Score 257; DB 14; Length 741; Best Local Similarity 56.3%; Pred. No. 6e-22; Matches 49; Conservative 13; Mismatches 25; Indels
APPLICANT: ISHIXAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
TITLE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR FILING DATE: 2001-05-30
PRIOR FILING DATE: 2001-05-30
PRIOR FILING DATE: 2001-05-30
PRIOR FILING DATE: 2001-06-30
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 8100
LENGTH: 741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651 AAVGCSNVSLKVVPLTTPVPGADAYLE 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: May 11, 2004, 12:25:48
Job time : 11.385 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8100
```

ö

```
em htg other:*
                             htg hum: *
   em_vi:*
168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2 1/USPFO spool/USO9917376/runat 11052004 114535 28241/app query.fasta_1.1429
-O=/cgn2 1/USPFO spool/USO9917376/runat 11052004 114535 28241/app query.fasta_1.1429
-DB=GenEmbl -OFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIXE=50 -MINLEN=0 -MAXLEN=2000000000
-USRE-USO9917376 @GCN 1 1.7742 @runat_11052004 114535 28241 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NGG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO MMAP -LARGEQUERY -NGG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                                                                                       May 12, 2004, 03:40:05; Search time 1279.29 Seconds (without alignments) 3015.368 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                         470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6940544
                                    GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                              - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3470272 segs, 21671516995 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                           US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gb_ba:

gb_htg:

gb_on:

gb_on:

gb_ph:

gb_pr:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro:

gb_ro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_fun: *
em_hum: *
em_in: *
em_mu: *
em_om: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_sts:*
em_an:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_or:*
em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GenEmbl: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                 Run on:
```

em_htgo_other:* em_htg_mus:*
em_htg_pln:*
em_htg_pln:*
em_htg_mam:*
em_htg_vrt:*
em_htg_vrt:*
em_htgo_mus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ľ	ġ.
۲	4
£	Ľ
•	ς
3	Ε
5	E,
t	2

Description		0500	9036	0058	0025	05023	98814	337 Caldibac	04 Caldibac	88933	3614	2 Truncated	3 Truncated	42 Caldi	Truncated	029	737 Caldicel	Ē	$\overline{}$		38	AF036924 Caldicell	529	5 Bacillus	B.subtili	B.subtili	Bacillus	22	subtili	~	_ !	7	8		~ ·	32	DNA segue	5164	8170 B	602 Cē	6308 Caldocell	6588 Bac	084 Baci	363635 Bacil	5197	
SUMMAKIES		05	70003	9	22	AP005023	3	AF163837	AF200304	PAE488933	BSP133614	E35142		AF078038S5	E35100	AB008029	AF078737	E35099	AF078038S3	ATZ86105		CASR69XYN2	AF355629	BSEGLSG	BACGLUB	BS14GLUC	BSBGLUC2	AY044252	BC170DEGR	BSUB0010	BACCMCASE	AY183475	AF045482	CDCMANA	CDCCELA	BACCELD	E05425	AB016164	A28170	CSCELB	CSU16308	BACCELAA	SU2	963	519	
DB	1	9	ø	ø	ø	Н	Н	-	Н	Н	Н	9	9	~	φ	Н	Н	9	Н	ч	Н	Н	Н	Н	Н	Н	Н		Н.			-	н.	н,	-	-	9	Н	ø	Н	Н	Н	Н	Н	-	ŀ
Length		2289	8	9	9	7	S	9	m	0	9	N	2	0	6416	4743	11707	11707	787	5513	3262	5437	1523	1800	1920	2314	2435	2589	26170	207829	1928	1593	2084	4977	5439	2175	2175	2587	2977	4241	5284	2831	1553	2152	6555)
* Query Match		8.66	99.1	ω.	'n	4.	48.6	8	45.6	44.6		41.6			41.4	ä	•	٠	40.7	ö	•	40.5	40.4	40.4	40.4	40.4	40.4	ö	0		٠	30.8	 o	39.7	٠. ص	φ.	φ.	39.6	ω.	38.8	æ	38.4	38.3	۲.	36.7	
Score) !	469	466	462	436	257	28	25	14	209.5	9	95	95	94	94	92	92	92	91	191	90	0	190	190	190	190	190	190	190	190	188	187	۲,	186.5	86	186	186	Н	82	82	182.5	8	180	177	172.5	•
esult No.		7	0	ю	4	Ŋ	9	7	80	0	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	30	31	32	33	34	32	36	37	38	39	40	41	42	43	44	

ALIGNMENTS

```
1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGACACTGTGAC 1929
                                                                                                                                                                                                                       1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                      1930 TGGGCGGCGATGGGGGTGTGGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTCGATTTGTCGACGGTG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2605 GTGTCGGGTGGGGTGAAGGTGCAGTATAAAGAATAATGATTCGGCGCCCGGGTGATAATCAG 2664
                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                                                                                        61 TrpalaalametGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                      41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                     1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspAsrAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                       21 IleLysproGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Acidothermus cellulolyticus
Acidothermus cellulolyticus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ding,S.Y., Adney,W.S., Vinzant,T.B. and Himmel,M.E.
Thermal tolerant avicelase from Acidothermus cellulolyticus
Patent: WO 03012090-A 2 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. .2869
/organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2869
2 86
0 0
0 0
                                                3687
87
1
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Frankineae, Acidothermaceae, Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AX700058 2869 bp
Sequence 2 from Patent WO03012090.
AX700058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1990 CCGACGCGGACACCTACCTGCAG 2013
                                                                                                                                                                     US-09-917-376-4 (1-89) x AX700036 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-4 (1-89) x AX700058 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AX700058.1 GI:29536021
                                              9.53e-48
466.00
100.00%
98.86%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.19e-47
462.00
100.00%
97.73%
98.30%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                     Best Local Similarity:
                                                                                    Percent Similarity:
                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
AX700058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1363 GTGTCGGGTGGGGTGAAGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1423 ATCAAACCGGGTCTCCAGTTGGTGAATACGGGGTCGTCGTCGGTGGATTTGTCGACGGTG 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1543 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ding,S.Y., Adney,W.S., Vinzant,T.B., Himmel,M.E. and Decker,S.R. Thermal toblerant callulase from Acidothermus cellulolyticus Patent: WO 03012109-A 2 13-FEB-2003; Midwest Research Institute (US)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                Acidothermus cellulolyticus
Acidothermus cellulolyticus
Bacteria, Actinobacteria, Actinobacteridae, Actinomycetales,
                                                                                                                                                                                 Ding, S.Y., Adney, W.S., Vinzant, T.B. and Himmel, M.E.
Thermal tolerant mannanese from Acidothermus cellulolyticus
Patent: WO 033012110-A 2 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                    linear
                                                                                                                                                                                                                                                                                 /organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                     228
88
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                     Frankineae; Acidothermaceae; Acidothermus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. 3687
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="texon:32630"
/note="Segment of GuxA"
                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
             Sequence 2 from Patent W003012110.
AX700050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AX700036 3687 bp
Sequence 2 from Patent W003012109.
AX700036 1 GI:29536019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-4 (1-89) x AX700050 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                               AX700050.1 GI:29536020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           synthetic construct
synthetic construct
artificial sequences.
                                                                                                                                                                                                                                                                                                                                                                                     2.21e-48
469.00
100.00%
100.00%
99.79%
                                                                                                                                                                                                                                                                  1, .2289
                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 2
AX700036
LOCUS
DEFINITION
AX700050
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                   ORGANISM
                                                ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                     ORIGIN
```

ò 엄 ò g ò g ò g ò 셤

g

g ð 셤

ò

```
Submitted (29-MAR-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:blo@mite.go.jp, URL:http://www.bio.nite.go.jp/, Tel:81-3-3481-1933, Fax:81-3-3481-8424)
This work was done in collaboration with Haruo Ikeda (*1), Jun Ishikawa (*2), Akiharu Hanamoto(*3), Chiqusa Takahashi (*3), Mayumi Shinose(*3), Hiroshi Horikawa (*4), Hidekazu Nakazawa (*4), Tomomi Osonoe(*4), Norihizo Kushida (*4), Hidekazu Nakazawa (*4), Tomomi Osonoe(*4), Norihizo Kushida (*4), Hideshi Kikuchi (*1,*7) and Satoshi Omura (*1,*3) and Satoshi Omura (*1,*3) and Institute for Life Sciences, Kitasato University *1 Kitasato Institute of Infectious Diseases and Justional Institute of Technology and Evaluation *5 School of Science (Technology and Evaluation *5 School of Science (Technology and Evaluation *5 School of Science (Office (Inversity of Tokyo *7 RIKEN, Genomic Sciences Center *1 Kitasato University of Tokyo *7 KIKEN, Genomic Sciences Center Following url is also available.

**Putter,/avermitilis.ls.kitasato-u.ac.jp.**

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers*

**Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /trānslation="MAGPPRSATRSRTTADREDSAHPAPDRADEPGYRRQIKVQANLQ
EGHHALARRAGQLYQRYQDGMEDQIGALGLVLNALVLFNTRYMDAALTQLRADGFEVR
DEDVARLSPFMRHHINVLGRYSFQLPDLE"
                                                                                                                                                                                                                                                                                                                                                                                                                               Ikeda,H., Ishikawa,J., Hanamoto,A., Shinose,M., Kikuchi,H.,
Shiba,T., Sakaki,Y., Hattori,M. and Omura,S.
Complete genome sequence and comparative analysis of the industrial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /noTe="This strain is also named as strain: ATCC 31267, NCIMB 12804 or NRRL 8165.~synonym: Streptomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 (bases 1 to 299175)
Omura,S., Ikeda,H., Ishikawa,J., Hanamoto,A., Takahashi,C.,
Shinose,M., Takahashi,Y., Horikawa,H., Nakazawa,H., Osonoe,T.,
Kushida,N., Director-General of Biotechnology Center, Shiba,T.
Sakaki,Y. and Hattori,M.
                                                                                                                                                           Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y. and Hattori, M. Genoe, T. Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary
                        Streptomyces avermitilis MA-4680
Streptomyces avermitilis MA-4680
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                        Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001)
21477403
11572948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Streptomyces avermitilis MA-4680"
/mol_type="genomic DNA"
                                                                                                              Streptomycineae; Streptomycetaceae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'product="putative transposase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             complement (474. .740)
/note="IS630 family, truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             microorganism Streptomyces avermitilis
Nat. Biotechnol. 21 (5), 526-531 (2003)
22608306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            db_xref="taxon:227882"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /protein_id="BAC68155.1
/db_xref="GI:29604084"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     complement (155. .550)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       In4561, truncated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /strain="MA-4680"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                codon start=1/trans1 table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            avermectinius"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           note="SAV445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . .299175
                                                                                                                                                                                                                                                                                                     metabolites
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12692562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                    MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
                                                                                                                                    REFERENCE
                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE
                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                              TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AP005023 299175 bp DNA linear BCT 10-MAY-2003 Streptomyces avermitilis genomic DNA, complete genome, section
                                                                                                         PAT 03-APR-2003
                                                          2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292 GGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACGCCGACGCGGCGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84
     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            232 TGGTTCACCCGGGGATGGTGGGTCGTCGACACTGGTACAACTGTGACTGGGCGGCGATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65 GlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAsp
  ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Thermal tolerant exoglucanase from Acidothermus cellulolyticus Patent: WO 03012095-A 2 13-FEB-2003; Midwest Research Institute (US) Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArgTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Decker, S.R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Adney, W.S., Ding, S.Y., Vinzant, T.B., Himmel, M.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3365
81
1
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Frankineae; Acidothermaceae; Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                                                        3365 bp
Sequence 2 from Patent WO03012095.
AX700025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                         2845 CCGACGCGGACACCTACCTGCAG 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4 (1-89) x AX700025 (1-3365)
                                                                                                                                                                                                                        88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cellulolyticus
                                                                                                                                                                                                                        ProThrAlaAspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AX700025.1 GI:29536018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AP005023 BA000030
AP005023.1 GI:29604083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.72e-44
436.00
97.62%
96.43%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lantz Mccarter, S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .3365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Acidothermus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         352
        41
                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
AP005023
LOCUS
                                                                                                                                                                                                                                                                                                                                                          AX700025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                     LOCUS
```

셤

ઠે

g

ò

요

8

a

8

d

```
105461 GCCCGCTACTACTTCACCCGGGACAGCGGCTCGCCCACCGTGAACGCCTGGTGCGACTAC 105520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /trans._touse=1.1
/product="hypothetical protein"
/protein.id="BACGB165.1"
/db_xref="G1:29604094"
/translation="MNPYAPASGIEQRPGWSWLDEYLWCSCDIIADLAEGRELEQRPLV
ATVARLEAGERVLAVGPAQRVTWRALGDGRYMHONAVAFGSPAFVLGSMASMATSALG
NSARRRAAANDAQPRCGARTPGLHVLDVLDVHGTMLRWSWHLGTSSTRRPSSALRALL
                                                                                                                                                                                                                   /translation="MVTAVRARARAGSAWQACRGARRSSQHSSEGVRRMDEKEFLRTV
AERTGLSRQEAADLTWATLETLARRLSPGEARDLITELPEGLAEAVRRGTTDR1ERFG
HHDSVRRVAERTMLKEEEADRGVRTVLAVLREAISEKEFNDLMSQLGTDFSQALESAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVGARVRWQVDTWITAPILSVSEQADVSDIAEAVRSHLDGAEWDVAIFLTDFPRRARLYP
ISVEVDTALRAALISLPALGVRRLRRRVRQAVVDVVRELVAHEETPRHPEVIGRPPRE
ISVEVDTREPGPEOPORYVVPGLRCHTRLVAGMYANRPRILFGSLISRALAGVFATRYTLF
INSATWILATALGAVQHAVITAGSAVALTLMI VÜNDLUMERGTHLPARHHPLYPYNLV
TLVTTALAVVVLAALEATIVAVSFLLIDASVJRKFIGRQVGVGDYLFLELAWFITAIAM
IGGARGTSLEGDEKVSDAAYGRRQRERQRMLRQMPRPERGERRGPDTTDG"
COMPLEMENT (6986. .7498)
SAPRITDIEDPEVHRTLRRSLEDLGAPAHLRPAAPRTHGGRRRCPLVHGPVRPRLAAV
                                                                                                                                                                                                                                                                                                                                                                                                                    /product="hypothetical protein"
protein id="BAC68164.1"
/db_xref="GI:2964093"
/translation="MIMTADVPETSSGRSVRVALVADPGLPSE1ARALAPRLPDHLRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /translation="MATLLRRVCPHEGVRIAAVPAGFVLSRRPWTASVYRCAVVGPQS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EDRREVEGSRTPVASASSRLRPCRRRLSVVRCPRRAVPSQAVRSRFRPGTTRFLGDVL
VYQRHRDALHARVRHVI DDVHPGLGRGAARPVR I ATHSLGGVI AVDMATANVFLMTES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LVTFGSQATFFHQPVPGDLLCQMIAFRMRKHQEAETGALPPETGFIAHAGQGW"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             299175
49
13
25
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                           /product="hypothetical protein"
/protein_id="BAC68163.1"
/db_xref="G1:29604092"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /transl_table=11
/product="hypothetical protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /protein_id="BAC68166.1"
/db_xref="GI:29604095"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-4 (1-89) x AP005023 (1-299175)
                                                                                                                       table=11
                                                                                                                                                                                                                                                                                                                                                                    codon_start=1/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             table=11
                                                                                                                                                                                                                                                                                                                         5810. .6964
/note="SAV454"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7548. .8189
/note="SAV456"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               codon start=1/transl table=
                                                5183. .5665
/note="SAV453"
                                                                                                codon start=1
transl table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             note="SAV455"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
/transl_table=
                         LVHGTAAGPCTPR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2.91e-19
257.00
71.26%
56.32%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ELVDLHVIAL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54.68%
                                                                                                                                                                                                                                                                                                                         5810.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                CDS
                                                                                                                                                                                                                                                                                                                         G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon start=1
/transI_table=11
/producE="putative TetR-family transcriptional regulator"
/protein_id="RAC68158.1"
/db_xref="GI:29604087"
/translation="WPTSTRRARERANTRERIIEAALHVLETEGIAALTIRRIATDVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | codon start=1
| transl_table=11
| product="hypothetical protein"
| protein_id="BAC68159.1"
| db_xref="G1:23604088"
| translation="MIYHINVTMKATATPEQIEGVLESWRNGRSNPAIKSFVVGRN
| HGGDYTYGAVEVVEHLGGLEPXLTHPTYQTPH.GLHLVERLEIFDVSDDDPDLNAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /protein_id=Bac68160.1"

/db_xref="G1:29604089"

/db_xref="G1:29604089"

/tb_xref="G1:29604089"

VaadasDsdperleelstuigagsigtivarilaeggesvrivtrigagsgpehplvdr
VaadasDsperleelstagagviytganepsytvmerlipploytaaitaakandavlalt
GSLYAYGPOPGGRANBETPPMAATGHKCRLERRWWEQALAGGIRTVERGSDVIGKDAN
GIYSLLIKSALQKGKAAWITGHLDMPHTFTFNGDMAQALITLASEERAWGRPWHVPSA
PAVTIRELARRYAVAAGREPPIKLIQIPRSVTRTAGLIVPIAREMAEMDYQWYAPFHMD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /transl_table=11
/product="hypothetical protein"
/protein id="BAC68162.1"
/db_xref="G1:20604091"
/translation="MRAVKESRVRVVGEIAAPPTGEALELVLHLGESIRGVRVESSDA
TARPDGLDFLAELPARGEWHTTALVRPSVEGEEAGGLFHPGAEPGESGPARRLCAWRE
                                                                                                                                           /tränslation="MPDTQNAGPSTSVTAAAVSTAGSVYSMMAPTSPVMAPPSPGRRR
EPGSLGAVRAYDVIRMLSRDGRPTPLGDAITYYGGSRRLCTSCA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YSAPVVYQHFANKDALVLELVAHGHRLMLSEFSQAAQEPDTDRRMTRIASQYVRFAGE
HPHLFQVMNDPMVDADERRRVAEPVIDVLKELLITSWSAAHDVVLADFDQACEILWGTL
YGIASLSYLGSVGNDRARYLAEQALHTILLGWRTETPANGRPASS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /trānslation="MTDAFHIDVPHTDAALAVIALSGEFDIAAAPAVRARALELIANG
HPDLVADLSGVTFCDSSGLGALVGIWRCAKDADGSLTLAAIPDRLSRLLSVTGMDAFL
PAYSSADAALAARQGNRTTA"
                                                                                                                                                                                                                                                                                                                                                                    /trānslation="MVVVAVGCALTSVPTAGGVVAEPVCVRRLTDQEGGKLQQIVRRG
SASSVRFRRAMMLLASAGGNRVPVIAQLV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            transI_table=11
product="putative anti-anti-sigma factor"
protein_id="BAC68161.1"
db_xref="GI:29604090"
                                                                    product="putative transposase"
protein_id="BAC68156.1"
db_xref="G1:29604085"
                                                                                                                                                                                                                                                                 /transI_table=11
/product="putative transposase"
/protein_id="BAC68157.1"
/db_xref="GI:29604086"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /trans]_table=11
/product="hypothétical protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATETTETFGLTATDLDTAVRNEVG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              complement (4130. .4498)
                                                                                                                                                                                                                                                                                                                                                                                                                    complement (1954. .2571)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  complement (4130. .4498)
                                                table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="SAV451"
/codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    note="SAV449"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   note="SAV450"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         635. .5156
note="SAV452"
                                                                                                                                                                                                                                                                                                                                                                                                                                              note="SAV448"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            codon_start=1/transl_table=
                         codon start=1
trans1 table=
                                                                                                                                                                                                                     note="SAV447"
                                                                                                                                                                                                                                               codon start=1
transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              codon_start=1
transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /gene="bldG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene="bldg"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .4029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CDS
```

```
/translation="TPTPTPTPSAAPTPTPSAGGSLVVQYRAADTNAGDNQLKPHFRI
VNRCTSSVPLSELTIRYWYTVDGDKPQVFNCDWAQVGCSNVRGSFVKLSTGRTGADYY
IEITFTSGAGSLAAGASSGDIQVRINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 (bases 1 to 4567)
Sunna A., Gibbs, W. D., Chin, C.W., Nelson, P.J. and Bergquist, P.L.
A gene encoding a novel multidomain beta-1,4-mannanse from
Caldibacillus cellulovorans and action of the recombinant enzyme on
                                                                                                    1694 GCTCAAGTGGGCTGCGCCAACGTGAACGCGTCCTTCACGTCGTG-----TCGGCTCC 1747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Direct Submission
Submitted (29-UUN-1999) Biological Sciences, Macquarie University,
                                                                                                                                                           22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                    AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                SerglyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                         42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF163837

Acadibacillus cellulovorans multidomain beta-1,4-mannanase precursor (manA) gene, complete cds; and unknown genes.
AF163837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacteria, Firmicutes, Bacillales, Alicyclobacillaceae, Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers

1. .4567
/organism="Caldibacillus cellulovorans"
/mol_type="genomic DNA"
/db_xref="taxon:74586"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .. .60
'note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Appl. Environ. Microbiol. 66 (2), 664-670 (2000)
20120520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     North Ryde, New South Wales 2109, Australia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 (bases 1 to 4567)
Sunna, A., Gibbs, M.D. and Bergquist, P.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /protein_id="AAF22273.1"
/db_xref="G1:6651326"
                                          US-09-917-376-4 (1-89) x AY298814 (1-1957)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Caldibacillus cellulovorans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Caldibacillus cellulovorans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
/transl_table=11
/product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                      1748 GGCGCCGACACCTCCCTG 1765
                                                                                                                                                                                                                                                                                                                                                                                              82 ThralaAspThrTyrLeu 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF163837.1 GI:6651325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          note="orf1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NGQLIWGVEP"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . 513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               kraft pulp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10653733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AF163837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                             ઠ
                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                ò
1 (bases 1 to 1957)
Posta, K., Beki, E., Kukolya, J. and Hornok, L.
Phylogenetic relationships of If cel5B, a new endoglucanase encoding gene from Thermobifida fusca
Unpublished
(bases 1 to 1957)
Posta, K., Beki, E., Kukolya, J. and Hornok, L.
Direct Submission
Submitted (14-MAY-2003) Agricultural Biotechnology and Microbiology, Sent Istvan University Godollo, Pater Karoly str.1., Godollo 2103, Hungary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HYYPWWYKGDITTEDFYTAWEWVTERYKNNDTIVAÄDIKNEPHGKANETPRAKWDGST
DINRFHVYESTAGKRILAINPNMALILCEGIBIYBYKOQOWASSTDGRDYYSTWWGGNLR
GVAAHPYDLGAHQDQLYYSPHDYGPSYPEQPWFEGEWNRQTLTEDVWRPWNLYHEDD
IAPLIJGEWGGSPLGGONEKWWTALRSLIIDEKWHHFFWALNPNSGDYGGLLMYDWT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WDBAKYAFLKPALWQDANGKFVGLDHDVPLGGVGSTTGVSLNQYYGGGGPSQPPTEPT
PEPTEPTEPPEPTEPTEPPATGALEVYYRNNSLTSLADBSQ1PFGCHLLVNTGSSTVÜLAD
VETHYYFTNEPGGTLQPFTCDWAQVGGANVNASFTSLABAPGADTSLVLTLSGSLAPGAS
TELQCRIHTANWANPESDDXSRGTNYTDWELSEVITAVLGGTLVWGTPPA.
                                                                                                                                                                               AY298814 1957 bp DNA linear BCT 23-JUN-2003
Thermobifida fusca strain TM51 endoglucanase (ce15B) gene, complete
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RIVDSAGNEVWLTGANWFGFNTSERMFHGLWAANIEDITSAMAERGINMVRVPISTQL
LLEWKNGQAGPSGVNEYVNPELAGMNTLEVFDYWLQLCEEYGLKVMLDVHSAEADNSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     translation="MTPLTRRLRAGAAAIAIGASALIPLTSSPAAASGTADWLHTDGN
                                                                                                                                                                                                                                                                                               Thermobifida fusca
Thermobifida fusca
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptosporangineae; Nocardiopsaceae; Thermobifida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="ce15B"
/note="Region: cellulose binding domain; CBM-3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'note="Region: catalytic domain; GH5 family'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1957
45
16
22
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1. .1957
/organism="Thermobifida fusca"
/mol type="genomic DNA"
/strain="TMS1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /codon_start=1
/transl_table=11
/product="endoglucanase"
/protoin_id="AAP56348.1"
/db_xref="G1:31745733"
                                                                                    105581 GGAGCCGACCCTACCTCGAA 105601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                db_xref="taxon:2021"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bound moiety="CelR".07. .1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                              82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note="putative"
                                                                                                                                                                                                                                         AY298814
AY298814.1 GI:31745732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene="cel5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene="cel5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gene="cel5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.89e-18
228.50
70.93%
52.33%
48.62%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                        RESULT 6
AY298814
LOCUS
                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                  유
                                                            ò
```

```
Caldibacillus cellulovorans putative type IIIb cellulose-binding domain gene, partial cds; and beta-1,4-xylanase XynA precursor AP200304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /protein_id="AAF61648.1"
| Da xreeF== (31:7385019"
| Aranslation="WOGGSNLGSEVKLSTGRTGADYXIETTFTSGAGSLAPGASSGD
| IQVRINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYRNGQLIWGVEP"
                                                                                               178 CGGTACTGGTACACGGTGGAC---GGGGACAAGCCGCAGGTGTTCAACTGCGACTGGGCG 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 (bases 1 to 3237)
Sunna,A., Gibbs,M.D. and Bergquist,P.L.
Direct Submission
Submitted (29-07-1999) Biological Sciences, Macquarie University,
North Ryde, Sydney, New South Wales 2109, Australia
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                             235 CAGGTGGTTGTTCGAACGTGCGCGGCAGCTTCGTGAAGCTTTCGACGGCCGGACGGGC 294
                                                                                                                                                                                                                                43 ArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAla 62
                                                                                                                                                                                                                                                                                                           63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /evidence=not_experimental
/transl_table=11
/product="putative type IIIb cellulose-binding domain"
                                                                         3 GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys
                                                                                                                                                    23 ProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. .3237
/organism="Caldibacillus cellulovorans"
/mol_type="genomic DNA"
/db_xref="taxon:74586"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'gene="xynA"
'note="multidomain protein"
                                     US-09-917-376-4 (1-89) x AF163837 (1-4567)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Caldibacillus cellulovorans
Caldibacillus cellulovorans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
/transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                          295 GCGGACTACTATATCGAG 312
                                                                                                                                                                                                                                                                                                                                                                                    83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       codon start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AF200304.1 GI:7385018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'note="orf1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             384. .3149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11065373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  α
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF200304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                           g
                                                                                                                                                    ò
                                                                                                                                                                                        셤
                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /trānslation="MGASGVPSRTAQNGYGLKARCQKRRCVGRNLLGHRLFVVDDDBF
RFFAGARKRLGRFSRRQNRLGEAQNFESGTVTFVQPDDKRRIRRENIGQLFVRFQASI
GPKRLIEIAEQRQARARKQASDQPQLKNAVILHFIDQHMPYGVERMDPFHKFIEKQKS
                                                                                                                              /trānslation="MNRRLIARLSGMLAWVLIAAMLAYPKPEPAEAHGGWVFPATRT
XCYVDGKWHGGGDLMMINPACLDBALAISGNYQFWNWFGWLISNAGGRRRBIIPDGK
LCGPTASFDGWHGNGTDWWTTRLQPGATIVKNNAMAPHPGWYLYVTRDGWDPTQPL
KWSDLEPTPFSQVTAFNPINSGSPDGAEYSWQVQLPNKGGRHIIWNGSDSPBAFYN
                                                                                                                                                                                                          CSDVYFGSGPIÄYEFGDPREGGTMITPPPSĞTTPTPTPTFTPTFTPTFTPTVTVTV
TPTSPTPPSASGTLENEYRYGDTSATDNQMKPOLRIVNTGSQAVPLTELKVRYWTKN
STAGEVFCDWAQIGGSNITAAQFVSLSQPVSGADSYIELSFTGGSIPAGGNTGEIQNR
IHPTNWMNYNETDDWSYNGAQTTWGPSTRITLYRNGYLVWGTBPGGGSSPPTFVTPT
PTPTPTSTPTPTPTFTSTPTPSGGPNLSVNTQCLVGINHPHAWYRDRLSSSLQGIRSW
                                                                                                                                                                                                                                                                                                       GANAVR.I VLSNGCRWTKI PASEVADI I SQARTLGYRAVVLEVHDTTGYGEDAÄACSWT
YAVRWI ELEKULAGQENEY TVI OURBEYGNNYKONWVTDTRAKAVQALENAGINNTIN
VDARWIGOWS FTWRDNAFTI TWADPÇENLIVES I HWYGYYDTRAVOSY I ESFVINGEI
PLVI GEFGHMISDGDPNEQAI VQYAKQYNI GLFGWSWSGNGGGVEYLLDWYTNFNANSP
                                                                                                                                                                                                                                                                                                                                                                                  codon start=1
'transī_table=11
'product="multidomain beta-1,4-mannanase precursor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Jote="Region: beta-1,4-mannanase catalytic domain" evidence=experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            note="Region: cellulose-binding domain type IIIb"
evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      note="Region: cellulose-binding domain type IIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'product="multidomain beta-1,4-mannanase" 349. .1480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note="Region: proline-threonine linker"
evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          note="Region: proline-threonine linker"
evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4567
41
17
27
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene="manA"
evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein_id="AAF22275.1"
'db_xref="G1:6651328"
                                                                                           protein_id="AAF22274.1"
                                                                                                               db_xref="GI:6651327"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           product="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                            RVTLYRNGQLVWGVEP"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 codon start=1
trans1 table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1481. .1927
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  960. .3409
gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     719. .3409
'gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2033. .2884
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  928. .2032
gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        885. .2959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.3e-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 225.50
67.44%
47.67%
47.98%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
```

/product="beta-1,4-xylanase XynA precursor"

BCT 15-JUN-2002

```
/translation="MLKTAAKKSLTAMLAGTVMLTGYTGLWTGSQTAHASDQAIEIQA
DGINEARFLQLYAQLKDPANGYFSAEGIPYHSIETLLSEAPDYGHMSTSEAYSYMLWL
ETWYGHYTGDWSKLEAAWDNMEKYIIPVNEGDGNEEQPTMSKYNPNSPATYAAEKPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DQYPSNLANGQYAAGKDPLDAELKATYGNNQTYLMHWLIDVDNWYGYGNLLAPSHTSTY
YNTPORGOGSSVWBA I PHPSQDDKS FGRAGEGPWSLEYREGORSAGNARYAWATDADA
RAVQYLYWAKEMGYNNSEYLDSQRKKWGDFLRYGMYDRYFGYKGSAKNGTPTPGTGKDS
NWYLMAWYTSWGGGLGQGGDWAWRIGASHTHQAYQNPVAAYALSDPAGGLIPDSPTAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADWATLKRQLEFYTWLQSHEGAVGGGATNSIGGSYAAYPAGVSTFYDMAYQEAPVYR
DPDSNTWFGFQAWPLERVAEMYYILAESGDLTSENFQMAKKVITKWVDWTKDYVFVGE
RPVTDAQGYYLNAAGQRILGGTNAQVATTPAPGEFWIPGGGEWQGQPDKWNGFSTYTN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NPNFHAITKOPVODTGVLGSYIKALTFFPAAGTEAENGTLSÄKGGEÄKDLAGSLLDTAW
DYNDGVGIVTEEERKDYFRYFAKEIYLPANWSGTFGQGNTIPGTAGVPSDPAKGGNGV
YIGYSDLRPAIKQDPAWAYLDNKYKTSYNPTTKQWENGAPTFTYHRFWSQVDMATAYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GMAGDAQAVLNWTASTGAATYTVQRSTGTGAYANLATGLTALTYTDTTAVNGTAYNYR
YATRASGQTYRSNVYVLTPWAAPISTGTLEVQYRNGGSBSAGNAITPQFNLKNTGTTA
I DLSKVKVRYYFTKNSAADMSFWCDYAQLGSANVQGFSVAWRAKGTADTYLEISFSRS
GAGSLAAGABETGIIQTRFSKONWTNFPOQSNDYSFNAIQTAFGAWTKVTAYQDGAKVWG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SVATGVTGSTFTDTGLTNGTAYYYVITAVNSAGESAPSAQVTATPQAGVTVPGVFTLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EFDRLLGDSSTPGVQVPSAPAGVTAAGGDKQVVLSWNAAAGATSYSVKRATVNGGPYT
                                                                                                                                                                                                                                     Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                         Direct Submission
Submitted (31-MAY-2002) Diaz P., Microbiology, University of
Barcelona, Av. Diagonal, 645, 08028-Barcelona, SPAIN
Location/Qualifiers
                                                                                                                                                                                                                                                                         Sanchez,M.M., Pastor,F.I.J. and Diaz,P.
Paenibacillus sp. BP-23 family 48 cellulase. Cloning and
performance on cellulosic substrates
Unpublished
2 (bases 1 to 3509)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /gene="ce148C"
167. .3334
/gene="ce148C"
/product="cellulose 1,4-beta-cellobiosidase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         country="Spain:Catalonia, Ebro river delta"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              product="cellulose 1,4-beta-cellobiosidase"
                                            3509 bp DNA linear Paenibacillus sp. BP-23 cel48C gene for cellulose AJ488011
                                                                                                                                                                 ce148C gene; cellulose 1,4-beta-cellobiosidase.
Paenibacillus sp. BP-23
Paenibacillus sp. BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BP-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62. .3337
/gene="ce148C"
/Eq. number="3.2.1.91"
/function="cellulose hydrolysis"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .3509 /
/organism="Paenibacillus sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /mol_type="genomic DNA"
/strain="BP-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     db_xref="taxon:198119"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
/transl_table=
                                                                                                                                                    GI:21449823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20. .25
/gene="ce148C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             39. .45
/gene="ce148C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'gene="ce148C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20. .25
/gene="ce148C"
                                                                                                                                                    AJ488933.1
                                                                                                                                                                                                                                                                                                                                                                                            Diaz, P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -10_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           promoter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                                                                 DEFINITION
                      RESULT 9
PAE488933
LOCUS
                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                                                                               AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene
                                                                                                                             ACCESSION
                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                           DEQDRKLILKHFNSVTPGNVLKFDSTEPQEGVPNFTESDKAVAFAVQNGMKIRGHTLI
MHNOTPNWPVEVDSNGNLVSKENLYQYBREHIKTVSVRKKGITSTAMDVNTBVIDEQQD
GLRRSLWYQTAGEST TEKAFOFAHEADPNALLFINDYNTHESGKSQALYNLVORLKSK
GIPVHGVGHQTHINITWPSISEIENSLVKFSNLGVVQEITELDMSIYNNSGXYDTLD
SADAQQOATRYRQLFRNFTRSSLIGNVTFWGADDAYTMLKFPVVNDWPLLFDEQL
KAKRAYWAVUGTVDRSPTPTFTSTAPPTPTPTPTPTPTPTPTPTPTPTPTSTPTPASASGTL
RVEYRVGDSSATDNQMKPQLRIVNTGSQAVPLIELKVRYWYTKNSTQAEQYFCDMAQI
                                                                                                                                                                                                                                                                                                   WSYNGTQTTWGPSTRITLXRNGVLVWGTEBGGGSSTPTPTPTPTPTFTSTPTPTPTPT
STPPTPTPTSTPTPTPTPTPTPTPFPTPSAGGNIV VQXRAADTNAGDNOLKBHFRLVNR
GTTSVPLSELTIRWTVDGDKPQVFNCDWARVGCSNLAGSIVKLTTGRTGRDYTLEI
TFTSGAGSLAPGANSGDIQARINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYRNGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2871 TGGGTCGGATGTTCGAACCTGCGCGACGCGGCACGGGGG 2930
GCSNIRAQFVSLAQPVSGADSYIELSFTGGSVPAGGNTGEIQNRIHFTNWMNYNETDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2694 GGCAACCTGGTCGTGCAGTACCGCGGCGGCGAACGCGGGCGACAACCAGCTGAAG 2753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2754 CCGCATTTTCGGATTGTGAACCGGGGACGAGCGTACCGTTGTCGAGCTTACGATT 2813
                                                                                   SAWVKLPSGSSNTRISMTMQRTMQDTVYYEQIYFDTALSGNWIQLKAQYKLYEPAVNL
QVYPEAPDHATQSFYIDDVRIEQLPDLPKTVEENIPSLKDVFAGRFPIGTAFENFELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               note="Region: beta-1,4-xylanase catalytic domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="Region: proline-threonine linker"
2697. .3146
/gone="xynk"
/note="Region: cellulose-binding domain typeIIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             note="Region: cellulose-binding domain typeIIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 note="Region: thermostabilizing domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="Region: proline-threonine linker"
2076. .2519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3237
40
16
29
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 product="beta-1,4-xylanase XynA'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x AF200304 (1-3237)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2931 GCGGACTACTTGAG 2948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                             384. .482
/gene="xynA"
483. .3146
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     966. .1946
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     947. .2075
gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2520. .2696
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.99e-16
214.50
65.12%
46.51%
45.64%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       183. .965
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                            LVWGVEP"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                             sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                         mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
0
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

ઠ

ò

ò

ò

ò

```
(bases 1 to 2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                      sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 11
E35142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                               2939 ACGCCGCAATTTAATCTGAAAAATACAGGTACAAACGGCAATTGATCTGAGTAAGGTGAAG 2998
                                                                                                                                                                                                                                                                                                                                              2999 GTTCGTTATTATTTCACGAAGGAC---AGCGCAGCGACATGAGCTTCTGGTGTGACTAT 3055
                                                                                                                                                                                                                                                                                                                                                                                                      3056 GCACAGCTTGGCAGTGCCAATGTGCAGGCAGCTTTGTCGCTGTAAATCCGGCGAAGGGC 3115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BCT 09-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pastor, F.I., Pujol, X., Blanco, A., Vidal, T., Torres, A.L. and Diaz, P. Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its
                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                         LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                      62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                                                                                                                               ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                               SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Direct Submission
Submitted (10-MAR-1999) Diaz P., Microbiology, University of
Barcelona, Av.Diagonal 645, Barcelona-08028, SPAIN
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       performance in pulp refining
Appl. Microbiol. Biotechnol. 55 (1), 61-68 (2001)
21129642
                                                                                  3509
40
21
25
1
                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AJ133614.1 GI:4490765
calB gene; cellulase; glycosyl hydrolase.
Bacillus sp. BP-23
Bacillus sp. BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1. 4161
/organism="Bacillus sp. BP-23"
/mol type="genomic DNA"
/strain="BP-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                            Indels:
                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref="taxon:89769"
398. .3525
/gene="celB"
398. .404
/gene="celB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4161 bp
                                                                                                                                                                                   US-09-917-376-4 (1-89) x PAE488933 (1-3509)
/EC_number="3.2.1.91"
3456. .3472
/gene="cel48C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                               3116 ACGGCAGATACGTATCTGGAG 3136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bacillus sp. BP23 celB gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                              82 ThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                437. .442
/gene="celB"
514. .519
/gene="celB"
532. .3525
                                                                                 9.26e-16
209.50
70.11%
45.98%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 (bases 1 to 4161)
                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BSP133614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AJ133614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11234960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Diaz, P.
                                                                     Alignment Scores:
                terminator
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -35_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -10_signal
                                                                                                                                                                                                                N
                                                                                                                                                                                                                                                                       22
                                                                                                                                                                                                                                                                                                                               42
                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
BSP133614
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                         ORIGIN
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                  원
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
```

```
AWWGPAEVWOWNR PSFKI DASCPGSDLAAETAAALAASS I VFADSDPVYŠAKLICOHAK
ELYRFADTYRGKYTDCITDAAFYNOWGYBDELAWGGAWLITATNDNAYLISKALSAA
BLWSTSGGSANWP YTWTGGRDS KYAQAILLAR I TSNLIMPEATKET QSTERNLDYWT
VGTNGGRVKYTPGGLAWLDQWGSLRYAANAAFI SFVYSDWYSDPVKKSRYQNFATSOI
NYLLGDNPGKSSYVOYGYGONSPOHPHTAHSKSWMNREDI PANTHRILIYGAWYGSDNA
SDQYTDDIGDYVSNRYATDYNAGFTGALAKNNLLYGQNHQPLANFPAEVKGDBYFVB
AAVRSSGSNYTEI RALLANRSGWPARMGDQLSFKYFLDLSEVYAAGRTVSDVQVTVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEGATYSQEVUYDAARRIYATTANFSNTKIYEGGEGYKRKEVQFRITGEPQGAWNPAND
PSYQNLTTGNPUKSNYIPYDAGVKVSGQEPGVTPVAVPAAPAGVQAVAGNSQVALINW
SASAGANSYTVKRAEVSGEPYTTVAAGVNGITYNTGLINGETYYYYVYTAVNSAGESP
ASVQVSGWPQATTVPGATVLSGTAGNNQINILSYTAASGGASSYTVQRAVAGGTYTDVA
TGLAVLIYNDTTALIVOTSYSYRVAANASGGTLSNVWILTPSGEPPSTGTLEVQYRSG
GSGNSSNAVTPQFNLKNTGTQAIDLSTVKIRYYFTKDGTEELSFWCDYAQVSGANVQG
MFVANNPAKGTADTYVEISFTSGAGSLAAGAETGVIQTRFSKNNWSAFDQSNDYSYDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :::|||||||::||||||:::||||||:::||||||:::
3187 ATTCGTTATTACTTTACCAAAGAC---GGCACGAGGAGGAGCTGTCCTTCTGGTGTGATTAT 3243
                                                                                                                                                                                                                                                                     LQKAIYFYEAQRSGPLPANNRVEWRGDSGMQDGADVGVDLTGGWYDAGDHVKFGFPMA
ASATMLAMSVVEYADGYEQAGQLEEIKDNIRWATDYFMKAHTKPNELWGQVGAGNTDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 18-JUN-2001
                                                                                                                                                                                                                                              /translation="MKGSWWRRVAIVTLSAGLLAGSFSMNTGLRQADAAAGDHNYAEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3067 ACAGGTACGCTTGAGGTGCAGTATCGAAGCGGGGGTTCAGGCAATTCCAGCAATGCAGTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4161
38
21
27
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
Mismatches:
Indels:
Gaps:
'EC_number="3.2.1.4"
'standard_name="endoglucanase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                               /protein_id="CAB38941.1"
/db_xref="G1:4490766"
/db_xref="GOA:Q92411"
/db_xref="SPTREMBL:Q92411"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SKTAFAAWNKVTAYQGNTQVWGLEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Truncated cellulase composition.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x BSP133614 (1-4161)
                                                                                                            'product="cellulase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3304 ACAGCGGATACGTATGTGGAG 3324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             82 ThralaAspThrTyrLeuGln 88
                                                                            _table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               532. .633
/gene="celB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         E35142.1 GI:13018967
JP 1999221086-A/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.83e-15
204.50
67.82%
43.68%
43.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  synthetic construct
synthetic construct
artificial sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
```

```
238 AGTGGAGCGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-4 (1-89) x E35143 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2.6e-14
195.50
60.23%
43.18%
41.60%
                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                            Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10706665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 of 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                  SELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AF078038S5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SEGMENT
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                     MORELAND DAVID GIBUSU, HYU MORGAN, DAIAN PURATONIOTISU WILLIAM C12N15/09, C11D3/386, C12N1/21, C12N9/42//(C12N1/21, C12R1:19), PC (C12N9/42, C12R1:19), C12N15/00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAT 18-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 AAGATAAGATACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :::
1 ATGGGAAGTGGTGTGAAGGTACTGTACAAGAACAATGAGAACAAGTGCGAGCACAGGTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 2029)
Paiji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G.,
Hyu,M. and Daian,P.W.
         Gurahamu, K.F., Moreland, D.G.,
                                                                                                                                                                                                                                          M DANIEL,
                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1. .2029
/organism='Artificial Sequence'
Pajji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Mo
Hyu,M. and Daian,P.W.
Truncated callulase composition
Patent: JP 1999221086-A 44 17-AUG-1999;
CLARIANT INTERNATIONAL LID
OS Artificial Sequence
PN JP 199221086-A/44
PP 21-SEP-1999
PP 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999
PF 21-SEP-1999

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2029
38
15
14
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:

    .2029
    ^organise="synthetic construct"
|mol_type="genomic DNA"
|db_xref="taxon:32630"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Truncated cellulase composition
Patent: JP 1999221086-A 45 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
OS Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2029 bp
Truncated cellulase composition.
E35143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGTGGAGCGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-4 (1-89) x E35142 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E35143.1 GI:13018968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.6e-14
195.50
60.23%
43.18%
41.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JP 1999221086-A/45.
unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unidentified unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                            SELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
No.:
                                                     TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
E35143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
```

```
Caldicellulosiruptor sp. Tok7B.1 glycosyl hydrolase 6 gene, partial cds; and CelE gene, complete cds.
AF078042.
                                                                                                                                              MORELAND DAVID GIBUSU, HYU MORGAN, DAIAN PURATONIOTISU WILLIAM C12N15/09,C11D3/386,C12N1/21,C12N9/42//(C12N1/21,C12R1:19), PC (C12N9/42,C12R1:19),C12N15/00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |||||| :::|||
178 TGGGCACAGATAGGGGCAAGCAATGTGACATTTTTGTGAAGCTTAGCAGCGGAGTG 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 AAGATAAGATACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :::
1 ATGGGAAGTGGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the
                                                                                                    M DANIEL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Caldicellulosiruptor sp. Tok7B.1
Caldicellulosiruptor sp. Tok7B.1
Bacteria, Firmicutes, Clostridia, Clostridiales,
Syntrophomonadaceae, Caldicellulosiruptor.

1 (bases 1 to 6005)
Caldicellulosiruptor.

1 (bases 2 to 6005)
Milliams, D.P. Reeves R.A., Farrington, G.K., Anderson, P.,
Williams, D.P. and Bergquist, P.L.
Multidomain and multifunctional glycosyl hydrolases from extreme thermophile Caldicellulosiruptor isolate Tok7B.1

Curr. Microbiol. 40 (5), 333-340 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 (bases 1 to 6005)
Gibbs, M.D., Reeves, R.A., Farrington, G.K., Anderson, P.,
PN JP 1999221086-A/45
PD 17-AUG-1999
PP 21-SEP-1999 JP 1998283606
PR 19-SEP-1997 US 08/932571
PJ PALJI ANDERSON, PETAER L BAGUKUISUTO, ROY GURAHANU K PARINTON, PI MORELAND DAVID GIBUSU, HYU MORGAN, DAIAN PURJ PC C12N15/09, C11D3/336, C12N1/21, C12N9/42// (C1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2029
38
15
115
1
                                                                                                                                                                                                                                                                                             /organism='Unidentified'
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                           Location/Qualifiers
1. .2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                             1. .2029
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
```

```
source
                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                   misc_feature
                                                                                          misc_feature
                                                                                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                   . No. :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMMENT
                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E35100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DLEEPADTTRSDAGYTAATGFYTSGGFIDDLGWAAWULYIATDRSUSCHTRABELMSEN
ANGTNIWTOCODDWYGTLIMLAKITGKELYKGAVERNLDHWTDR.TYTPKGMATITG
WGSLTAATTAATGACTLAATGFYTSGGFIDDLGWAAWULYIATDRTYTPTPKGMATITG
WGSLARAACWYADWGGCDSNKKTKYLNPAKSQIDYALGSTGRSFVVGFGTNY
POHPHRNAHSSWANSKIPPRILINGALVGGPGSBDSYNDDITDYVQNBYACDYN
AGIVGALAKWYQLYGGEPIDDPKAIETPTNDBIFYSEKGSNSQGPRYTBYISYISYI
GWPPRYTDKLSFWYFIDLLOAGSPBDVWVDTYIEGGKISGPYWDRINRYY
INDESCYKIYPGGGEBYEHKQAOFKISVPROPTPTPTPTPTPTPTTAATAAY
NNNLWGLEPGAATSTPAPTSPTPTPTPTPTPTPTPTPTTAATAAY
NNNLWGLEPGAATSTPAPTSPTPTPTPTTATATPTTTATAAY
NNTENPTKLSSGVSGADYYLEVGESSGAGOLQPGKOTGGB
NYTFNFYKLSSGVSGADYYLEVGESSGAGOLQPGKOTGFIGAS
NYTFNFYKLSSGVSGADYYLEVGESSGAGOLQPGKOTGFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPTVSATPTPAPTASPVGGSYWTPSESYGALKVWYANGNLSSPTNVLNEKIKIENVGT
APOLSRWYKYKYWYTJIOGEATOSVSYASSINPAPTIOVKEKKGANGGGADYVYBIGFK
SGAGYLAAGOSTKEI TRISIOKGSGSYNGSNINSSYREANSY IENBEKYTGYIDDGUFWE
BPGRNAQIKWWYANGNLGSWINVLNPKIKIENNGTTAVDLSRVKRYWYTIDGEATOS
SVYSSEINPAYTIOKEWYLAANGAGADYYVBETGEKSGAGVLAAGOSTKEIKLSIOKGS
GSYNGSNDYSVRSATCYIENBEVYTGYIDGATYWEREPSKGYKEPAGGYTPTPTGT
PTPTTTPTPTTTPTTTTPTTTTTPTPAYTPDVKISIDTSKGKTEPGGYTPTPTTFTT
SINGNAYSATTLQWYWSINGSNGYWESSETTRESPRAYBETGYN
SIRGNAYSATTLQWYGYYAKGONGTVSESETTREPSPRATHENCHYNDNYN
SIRGNAYSATTLQWYGYYAKGONGTVSESETTREPSPRATHENCHYNDNYN
SIRGNAYSATTLQWYGYYAKGONGTVSESETTREPSPRATHENCHYNDNYN
SIRGNAYSATTLQWYGYYAKGONGTVSESETTREPSPRATHENCHYNDNYN
VYMDEPINYLINKYGRSSSATGIKGYILDNEPDLMFTTHPRIHPQKVTCSELINKSVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /trānslation="GQPDTWNGTYTGNPNLHVKIVDYGTDLGITASLANALLYYSAAT
KEYGVSDBAAKNLAKELLDRMWNLYRDDKGLSAPEKRGDYKRPFEQEVYIPAGWTGKM
PNGDVIKSGVKFIDIRSKYKQDPDWQKLVSAYNAGEAPEFRYHRFWAQCDIAIANATY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LAKVIKTLDPDAEI FGPASYGFVGYLTLQDAPDMNQVKGNHRWFLSWYLBQMKKASDS
FGRRLLDVLDTHWYPBAQVGGYRICPOEDSYSFRDVA LAKMAGARTLWDPTYKYTQKG
QITAGENWINQWPEYLPLLPNIKADIDKYYPGTKLAITEFDYGGKDHISGGIALAD
VLGIFGEKYGVYMAARMGDSGSYAQAAYNIYLNYDGKGSRYGSTCVSAETTDVERMPVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LQKTIMFYEFQMSGKLPSWVRNNWRGDSGLDDGKDVGLDLTGGWHDAGDHVKFNLPMS
YSASMLGWAVYEYKDAFVKSKQLEHILNQIEWANDYFVKCHPSKYVYYYQVGDPTVDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |protein_id="AAK06394.1"
|db_xref="GI:12743885"
|translation="MQEMKAIKRVVSITALLVJTLSLCFPGIMPVKAYAGGTYNYGEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NFWGPAEVMQMKRPAYKCDLSNPASSVVAETAASLAVASVVIKERNSQKAASYLQHAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ASIEGEDDSTVHIILINRNYDRKLKAEIKMNNTRVYTGGEIYGFDSTSSQIRKMGVLS
Williams,D.P. and Bergquist,P.L.
Direct Submission
Submitted (27-OCT-1999) Biological Sciences, Macquarie University,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <1. .503
/note="Region: family 48 glycosyl hydrolase domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4296. .5869
/note="Region: family 44 glycosyl hydrolase domain"
2537. .2626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          note = "duplicated type IIIb cellulose binding site"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         725. .2053
'note="Region: family 9 glycosyl hydrolase domain"
                                                                                                                                                                         organism="Caldicellulosiruptor sp. Tok7B.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2054. .2545
'note="type IIIc cellulose binding site"
2627. .3106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  note="type IIb cellulose binding site"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 note="multidomain glycosyl hydrolase"
                                                                                                                                                                                                                                                                                                                                                                                              product="glycosyl hydrolase 6"
protein id="AAK06393.1"
db xref="GI:12743884"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NIQNNTITIEVPNLTVYHIVLTSSK"
                                                                                                                                                                                                                                                                                                         note="ORF6; multidomain"
codon start=3
transI table=11
                                                                                                                                                                                                      /mol_type="genomic DNA"
/isolate="Tok7B.1"
                                                                                                                                                                                                                                                           db_xref="taxon:80339"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    codon_start=1
evidence=experimental
                                                                                       Sydney, NSW 2109, Australia
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        transl table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        product="CelE"
                                                                                                                                            . .6005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EILFGNQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                TITLE
JOURNAL
                                                                                                                                                                                                                                                                                        SGO
```

```
OS UNKIDOWN
PN JP 199221086-A/2
PD 17-AGG-1999
PP 13-SEP-1998 JP 1998283606
PR 19-SEP-1997 US 08/932571
PI FAIJT ANDERSON, PETAER L BAGUKUISUTO, ROY M DANIEL, PI GURAHAMU K PARINTON GIBUSU, HYU MORGAN, DAIAN PURATONIOTISU WILLIAM PC C12N15/09, C11D3/386, C12N1/21, C12N9/42//(C12N1/21, C12R1:19), PC (C12N9/42, C12R1:19), C12N15/00
                                                                                                                                                                                                                                                                                                                                                             2648 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAAGGCCG 2707
                                                                                                                                                                                                                                                                                                                                                                                                                                        2708 TGGTTTAAGATAGTGAATGAGGCAGCAGCAGTGTTGATCTTAGCAGGGTTAAGATAAGA 2767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAT 18-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2768 rácrigéracacagragargargacaagccacagagrgcggra---rérgacrégégacag 2824
                                                                                                                                                                                                                                                                                                                                                                                                                  43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                      4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ., Roy, M.D., Gurahamu, K.F., Moreland, D.G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                     6005
38
14
32
1
                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ce 1. .6416
/organism='Unknown'.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Truncated cellulase composition
Patent: JP 1999221086-A 2 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. .6416
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6416 bp
Truncated cellulase composition.
E35100
                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x AF078038S5 (1-6005)
                    3107. .3223
/note="PT-linker"
4196. .4295
/note="PT-linker"
/note="PT-linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E35100.1 GI:13018925
JP 1999221086-A/2.
unidentified
unidentified
unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2885 GATTATTACCTGGAG 2899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 (bases 1 to 6416)
Paiji, A., Petaer, L.B.,
Hyu, M. and Daian, P.W.
                                                                                                                                          1.36e-13
194.50
61.18%
44.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84 AspThrTyrLeuGln 88
```

ò

8

ò 셤 ò ò

```
SNNMTIRGHTLVWHQQTPDWFFKDSNGNTLSKDALLSRLKQYIYDVVGRYKGKVYAWD
VVNBAIDBGQSDGYRRSNWYNICGPBYIEKAFIWAHEADPNAKLFYNDYNTENSQKRQ
FIYNLIKSLKEKGVPIHGVGLQCHINLDWPSISBIBNTIKLFKSIPGIBIHITELDMS
FYNDLKSSNYAVAPPRELLIKQQRYKELFDLFKKYNVYTNVTFWGLKDDSDSHLSKNY
GKKOYPLLPDRYNYRKXAFWSLIDDBIVPTTTLPTPPBIQLPTTPTPTPASYTPTPT
VTFTPTVTTTTPTPTPTPTPTPTPTNTRIR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             product="beta-glucanase"
protein id="BAA22939.1"
/db_xref="GI:2564015"
/translation="MEKKLILKHFNSITPENEIEPSELLINENTYNFSKADSFVNFAI
                                                                                                                                                                                                                                                                           2785 TACTGGTACACAGAGGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2841
                                                                                                                                                                                                                                                                                                                                                                                                                                          BCT 22-0CT-1997
                                                                                                                                                                                                                                                        43
                                                                                                                                                                                                                                                                                                                                TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                      83
                                                                                                                                                                                  23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AB008029 4743 bp DNA linear BCT 22-OCT-19
Thermophilic anaerobe NA10 gene for beta-glucanase, complete cds.
AB008029
                                                                                                                                                                          GlyValLygValGlnTyrLysAsnAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Direct Submission
Submitted (09-OCT-1997) Katsuhide Miyake, Nagoya University,
Department of Biotechnology, School of Engineering,; Furo-cho,
Chikusa-ku, Nagoya, Aichi 464-01, Japan
(B-mail:miyake@proc.nubio.nagoya-u.ac.jp, Tel:+81-052-789-4278,
Fax:+81-052-789-3221)
                                                                                                                                                                                                                                                      24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                      MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Miyake,K., Machida,Y., Hattori,K. and Iijima,S.
Characterization of a multi-domain cellulase from an extremely
thermophilic anaerobe strain NA10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="thermophilic anaerobe NA10"
/mol_type="genomic DNA"
/strain="NA10"
             6416
38
14
32
1
             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (bases 1 to 4743)
ake, K., Machida, Y. and Iijima, S.
                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'db_xref="taxon:67756"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                          x E35100 (1-6416)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   thermophilic anaerobe NA10
thermophilic anaerobe NA10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             codon_start=1
transl_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||| ||||||:::
GATTATTACCTGGAG 2916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AB008029.1 GI:2564014
           1.47e-13
194.50
61.18%
44.71%
41.38%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .4743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  beta-glucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unpublished
                                                 Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                              US-09-917-376-4 (1-89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (sites)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Miyake, K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria
 Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2902
                                                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                                44
                                                                                                                                                                                                                                                                                                                                                                                                        64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 15
AB008029
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
Alignment S
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                    Score:
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
21
                                                                                                                                                                                                                                                                 41
                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                     42 ValargTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                           AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                           SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                               22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                        US-09-917-376-4 (1-89) x AB008029 (1-4743)
                                                                                                                                                                                                                                                                                                                                                                                                                                Search completed: May 12, 2004, 09:26:13
Job time : 1315.29 secs
                                                                                                                                                                                                                                                                                                                                                                                                   1686 GGAGCGACTATTACTTGGAG 1706
                                                                                                                                                                                                                                                                                                                                                                                 82 ThralaAspThrTyrLeuGln 88
                                                                                                                                     1.79e-13
192.50
63.22$
43.68$
                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                              Scores:
                                                                                                                                                                                                                            N
                                                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                            Query Match:
                                                                                                                                       Pred. No.:
                                                                                                                              Alignment
                                                                                                           ORIGIN
                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                             원
                                                                                                                                                                                                                                                                 ሯ
                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                δ
```

Endol gen DNA encod DNA encod

NK-1 cell

Bacillus

thermo Nucleotid Nucleotid

Bacillus

Scoring table:

Searched:

Perfect score:

Sequence:

protein

Run on:

Pectate]

Prokaryot Fusion pr Fusion pr

Endo3 gen Cellulose C. cellul Clostridi Clostridi

DNA seque

Clostridi Clostridi

Chimeric

coli

```
mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
feed; paper pulp; biofuel; mannase; gene; ss.
Aax55662
Aad26526
Aad26525
Aal41028
Aal41028
Aal41028
Aag49820
Aaq49820
Aaq49820
Aaq4336
Aax90978
Aax90978
Aax75060
Aax7059
Aax7059
Aay27429
Aay27429
Aay27429
Aay27429
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
Aay27473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Aaf86248
Abk52403
Aad11046
Aad11044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Aad11045
Aad11043
Aaz27550
Aad22201
Aaf12881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Aaz92616
Ada02720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nucleotide sequence of the ManA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
1. .2289
/*tag= a
/product= "ManA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX24930
AAD11042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABK52403
AAD11046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAD11045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACA25846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAD11044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADA02720
                                                                                                                                                                                                                                                            37606XAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABZ77633 standard; DNA; 2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MIDE ) MIDWEST RES INST.
                                                                                                   1434
11488
22510
221710
221710
11488
1482
11314
11482
11303
11482
11303
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
11440
1140
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
11400
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adney WS,
     WO2003012110-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-FEB-2003,
  194.5
194.5
194.5
192.5
188
188
186
182.5
165
165
165
165
165
165
165
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
1150.5
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABZ77633;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ManA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT

        444446
        44446
        44446
        44446
        44446
        44446
        44446
        44466
        44466
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        44666
        446666
        44666
        446666
        446666

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acidother
A. cellul
Portion o
DNA seque
Active ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nucleotid
Nucleotid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nucleotid
                                                                                                                                                                                                       (without alignments)
3062.577 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Description
                                                                                                                                                                             ; Search time 123.455 Seconds
                                                                                                                                                                                                                                                                                                              470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Abz77634 1
Abz77632 1
Add22922 2
Abz76162 2
Aaq15178 1
Aax55660 1
Aad26568 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ab277633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6747726
                            GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd
                                                                                                                                using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3373863 seqs, 2124099041 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABZ77633
ABZ77634
ABZ77632
ADD22922
ABZ76162
AAQ15178
AAX55660
                                                                                                                                                                                                                                                                                                                                                                                BLOSUM62

: Xgapop 10.0 , Xgapext

Ygapop 10.0 , Ygapext

Fgapop 6.0 , Fgapext

Delop 6.0 , Delext
                                                                                                                                                                                May 12, 2004, 03:07:10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           geneseqn1980s:*
geneseqn1990s:*
geneseqn2000s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           geneseqn2003bs:*
geneseqn2003cs:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            geneseqn2001as:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 geneseqn2001bs:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             geneseqn2004s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      geneseqn2003as:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Geneseq 29Jan04:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            geneseqn2002s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match Length DB
                                                                                                                                                                                                                                                                                      US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3687
2869
3369
2000
2009
                                                                                                                                nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Command line parameters:
```

Database :

466 462 462 462 231.5 195.5

Score

Result Š

Genomic D Human NFA

Gaussia s

Clostrid Gaussia Asperqil]

Prokaryot

```
13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABZ77632;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABZ77632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RXXXXEX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1483 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TGGGCGCGCGATGGGCTGTGGGAATATCCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                                                                                                                                                                                                                                                                                        1363 GTGTCGGGTGGGGTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGCGGTGATAACCAG 1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATCAAACCGGGTCTCCAGTTGGTGAATACGGGGTCGTCGTCGGTGGATTTGTCGACGGTG 1482
                                                                                                           The present sequence encodes ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starting material to simpler carbohydrate units ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA
                                Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cellulose; sugar; biofuel; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrValArgTyrTrpPheThrArgAspG1yG1ySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                               Sequence 2289 BP; 463 A; 700 C; 688 G; 438 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                   2289
88
0
0
0
                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GuxA; cellulase, glycoside hydrolase, enzyme; detergent; pulp processing; paper processing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleotide sequence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCGACGCGGACACCTACCTGCAG 1626
                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-4 (1-89) x ABZ77633 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1. .3687
                                                                                     Example 1; Page 23; 46pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /*tag= a
/product= "GuxA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABZ77634 standard; DNA; 3687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                     4.13e-48
                                                                                                                                                                                                                                                                                                            469.00
100.00%
100.00%
99.79%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
WPI; 2003-248182/24.
P-PSDB; ABP73022.
                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                       polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABZ77634;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene; ss
                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABZ77634
 8
                                                                                                                                                                                                                                                                                                                                                                                                                                      임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
```

```
1750 GTGTCGGGTGGCTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1870 Accerceceracrestreaccececaresresecercerceacacreseratacaacresea 1929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence encodes a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1810 ATCAAACCGGGTCTCCAGTTGATGAATACCGGGTCGTCGTCGGTGGATTTGTCGACGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21 IleLygProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 3687 BP; 725 A; 1173 C; 1171 G; 618 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                         Decker SR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                      Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-4 (1-89) x ABZ77634 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Example 1; Page 23-24; 47pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1990 CCGACGCCGACACCTACCTGCAG
                                                                                                                                                                                                                                                                                                                      TB,
                                                                                                                          28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABZ77632 standard; DNA; 2869
                                                                                                                                                                                        28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.79e-47
466.00
100.00%
98.86%
99.15%
                                                                                                                                                                                                                                                                                                                      Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             03-JUN-2003 (first entry)
                                                                                                                                                                                                                                                       (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-239526/23.
P-PSDB; ABP73029.
                                                                                                                                                                                                                                                                                                                      Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
WO2003012109-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
```

detergent;

Key

```
2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                               2785 TGGGCGCGCGATCGGGTGTGGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 2844
                                                                                                                                                                                                                                                                                                                                                                                                            AviIII; callulose reduction; agricultural biomass; municipal solid waste;
glycoside hydrolase; avicelase; ds; gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucleotide is useful for reducing reliables in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents DNA encoding the Acidothermus cellulolyticus
                                        TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus avicelase AviIII DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Himmel ME;
                                                                                                                      Claim 17; SEQ ID NO 2; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
1. .2868
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /product= "Avilli"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Vinzant TB,
                                                                                                                                                                                                                                                 ADD22922 standard; DNA; 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.09e-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              462.00
100.00%
97.73%
98.30%
                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-810853/76.
P-PSDB; ADD22921.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (VINZ/) VINZANT T B (HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DING S.
ADNEY W S.
VINZANT T B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US2003108988-A1.
                                                                                                                                                                                                                                                                                                                               15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-JUN-2003
                                                                                                                                                                                                                                                                                         ADD22922;
                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (DING/)
(ADNE/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ding
                                                                                                                                                                                                                                  ADD22922
                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                   셤
                                                                                                                        8
                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGGTGGATTTGTCCACGGTG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2605 GTGTCGGGTGGGGTGAAGAGTGCAGTATAAGAATAATGATTCGGCGCCGGGTGATAATCAG 2664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence encodes a thermostable avicelase polypeptide, designated AviIII. AviII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20
                                          Avicelase, AviIII, glycoside hydrolase, enzyme, cellulase, biofuel;
detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ThrValargTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21 IlelysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seguence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                                 Location/Qualifiers
1. .2869
/*tag= a /partial
/product= "AviIII"
/fransl except= (pos:2869, aa:Xaa)
/note= "Xaa is an unspecified residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2869
22
00
00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
    Nucleotide sequence of the avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89) x ABZ77632 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 3; Page 24; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001WO-US023818
                                                                                                                          Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09e-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           462.00
100.00%
97.73%
98.30%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (MIDE ) MIDWEST RES INST.
                                                                                      cellulose; gene; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2003-248177/24.
P-PSDB; ABP73015.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                    WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                            13-FEB-2003
```

cells.

host

Ding S,

2869 86 2 0 0

2665

41

Н

ò 셤 ò g

Query Match DB:

Score:

g

ò

셤

ò

임 ò a ò 셤

ò

```
polypeptide is useful in the degradation of cellulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates atternative cellulase enzymes capable of assisting in the commercial scale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Guxl cellulase encoding
                                                                                                                                                                                                                                                                                              231
                                                                                                                                                                                                                                                                                                                                                                                                                                                 292 GGGTGTGGGGATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACGCCGACGCGGAG
                                                                                                                                                                                                                                                                                                                                                                                                          291
                                                                                                                                                                                                                                                                                                                                                                                 64
                                                                                                                                                                                                                                                                                                                                                                                                                                  84
                                                                                                                                                                                                                                                                                 24
                                                                                                                                                                                                                                                                                                                                 44
                                                                                                                                                                                                                                                                                                                                               172 CTCCAGTTGGAATACCGGGTCGTCGTCGGTGGATTTGTCGACGGTGACGGTGCGGTAC
                                                                                                                                                                                                                                                                                                                                                                                              65 GlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArdTyr
                                                                                                                                                                                                                                                                                                                                                                                 45 TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Detergents; pharmaceuticals; deinking; carboxymethylcellulose; ss.
                                                                                                                            Sequence 3365 BP; 669 A; 1117 C; 1041 G; 538 T; 0 U; 0 Other;
                                                                                                                                                                  3365
81
1
2
0
0
                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Portion of pAEC-1 contg. cellulase AE-1 gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /*tag= a
/product= "cellulase AE-1"
                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                       US-09-917-376-4 (1-89) x ABZ76162 (1-3365)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
192. .2540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             멾.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAQ15178 standard; DNA; 2600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90JP-00045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       90JP-00045465
                                                                                                                                                                 8.56e-44
436.00
97.62%
96.43%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (revised)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (revised)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OJIP ) OJI PAPER CO
                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JP03251174-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-FEB-1990;
                                                                                                                                                      Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-AUG-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
17-MAR-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-NOV-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Aeromonas
                                                                                                                                                                                                                                                                                 ß
                                                                                                                                                                                                                                                                                                                                 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAQ15178;
                                                                                                                                                                                                                    Query Match
                                                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
AAQ15178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Key
                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88888888888
                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                           음
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                     ACGGIGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                                                                                                                                                                    2605 GTGTCGGGTGGGGTGAAGGTGCAGTATAAGAATAATGATTCGGCGCGCGGGTGATAATCAG 2664
                                                                                                     40
                                                                                                                                                      9
                                                                                                                                                                                                     TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                      20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to a thermal tolerant Guxl peptide from A. cellulolyticus. The Guxl exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type III, and a carbohydrate binding domain type III. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New thermal tolerant Gux1 peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                     ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                      ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                     IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                       oux1; thermal tolerant; exoglucanase; glycoside hydrolase; cellulose;
biofuel; detergent; pulp; paper processing; feed processing; textile;
cellulase; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Decker SR;
   0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WS, Ding S, Vinzant TB, Himmel MB, Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                          A. cellulolyticus Gux1 protein encoding DNA
     Gaps:
                                                                                                                                                                                                                                                     ProthrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 22-23; 44pp; English.
                            US-09-917-376-4 (1-89) x ADD22922 (1-2869)
                                                                                                                                                                                                                                                                                                               2725
                                                                                                                                                                                                                              2785
                                                                                                                              2665
                                                                                                                                                      41
                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                               2845
                                                                                                     21
                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                         RESULT
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ጵ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   日文化文品文品文章
                                                                                                                                                                                                                                                                                                                                                                                                                              2142 AAGCCTCATTTCAATATCGTGAACAAGGGGGCTGCGGCAGTGCCTTTAAGCGAGCTGAGT 2201
                                                                                                                                                                                                                                                                                                                                                                                             2202 CTGCGATATTATTTCACAGCTGAC---GGCAATGACCAACTGCAATACAATTGTGACTGG 2258
                                                                               The plasmid, pAEC 1, was prepd. ligating chromosomal DNA contg. the gene (obtd. from Aeromonas) into pUC18. The gene can be used to produce recombinant enzyme which is used for the effective utilis- ation of biomass resources and the mfr. of pharmaccuticals and foodstuffs, and also for the detergent and delinking of waste paper. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
                                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                     SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                               opt.
                                                                                                                                                                                                                                                                                                                               LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                      ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                  AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1, Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing; ss.
                              Cellulase AE-1 for e.g. mfr. of pharmaceuticals and foodstuffs - of pH when carboxy:methyl:cellulose is used as substrate.
                                                                                                                                                          Sequence 2600 BP; 687 A; 649 C; 700 G; 564 T; 0 U; 0 Other;
                                                                                                                                                                                                    1
1
1
1
1
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA sequence encoding truncated cellulase Cel E3/B5
                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                 US-09-917-376-4 (1-89) x AAQ15178 (1-2600)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2319 AACGCCGATACCTACTTGGAG 2339
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ВР.
                                                             Claim 3; Fig 2; 8pp; Japanese
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX55660 standard; DNA; 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98EP-00810919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-00932571
                                                                                                                                                                                        1.62e-18
231.50
66.67%
48.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
 WPI; 1991-373412/51.
                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
         P-PSDB; AAR15625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unidentified
                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                               22
                                                                                                                                                                                                                                                                                                                                                                                                                  62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX55660
                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAX55660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                     원
                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                  8
```

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4,5, Cel B1/2, Cel 1/2/3, Cel 6 or Cel B3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B4/S extends from amino acid A1011 to P1424 or K4425 or N1426, and Cel B4/S extends from amino acid K635 to N1426 in the sequence shown in AAY13492; Cel B1 extends from amino acid X93 to D481, Cel E6 extends from amino acid Y133 to K1751 and the stability region extends from amino acid V133 to K1751 and the stability region extends from amino acid V13494. The new sequence shown in AAY13493; Cel E3/B5 is shown in AAY13494. The new carymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for.

processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-truncated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Active cellulase protein; alkalinophilic; textile processing; proteinase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TrpalaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                 cellulase proteins, useful in detergents and for producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :::
1 ATGGGAAGTGGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2029 BP; 650 A; 329 C; 520 G; 530 T; 0 U; 0 Other;
Farrington GK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2029
38
15
15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Active cellulase hybrid protein, E3/B5 DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
Daniels RM,
                                                                                                                                                                                                                                                    Disclosure; Page 41-42; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGTGGAGCGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            x AAX55660 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAD26568 standard; DNA; 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.53e-14
195.50
60.23*
43.18*
41.60*
Anderson P, Bergquist PL,
Morgan H, Williams DP;
                                                                                                                                                                   New truncated cellul
'stonewashed' denim.
                                                                                   WPI; 1999-315403/27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-4 (1-89)
                                                                                                                P-PSDB; AAY13494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAD26568;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAD26568
```

φ

```
The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the CelB5 amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing whits or for imparting stonewashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present sequence is E3/B5 hybrid protein DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ||||:::|||
| ATAAGGCCGTGGTTTAAGATAGTGAATGGAGGCAGCAGCAGCAGTGTTTAGCAGGGTT 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 AAGATAAGATACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ||||||| :::||| :::||| :::||| | ||:::
                                                                                                                                                                                                                                                                                                                                                                                                                    New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :::
ATGGGAAGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        detergent additive; stonewashed appearance; cotton-containing denim;
CelB5; thermophilic; commercial detergent; B3/B5 hybrid protein; ds.
                                                                                                                                                                                                                                                                                                                        Daniels R, Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                650 A; 329 C; 520 G; 530 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2029
38
15
14
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                              "E3/B5 hybrid protein"
                                                                                                                                                                                                                                                                                                                          Bergquist P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-917-376-4 (1-89) x AAD26568 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Col 71-74; 61pp; English.
                                                                               cocation/Qualifiers
                                                                                                                                                                                                                                                                                            LI
LI
                                                                                                                                                                                                                           98US-00136574.
                                                                                                                                                                                                                                                           97US-00932571
                                                                                                                                                                                                                                                                                         (CLRN ) CLARIANT FINANCE BVI
                                                                                                                                                                                                                                                                                                                        GK, Anderson P,
Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3.53e-14
195.50
60.23%
43.18%
41.60%
                                                                                                               /*tag= a
/product=
                                                                                                l. .1851
/*tag≃ ∂
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      elevated temperatures.
                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-081780/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAE16325.
                                                                                                                                                                                                                                                                                                                        Farrington GK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2029
                                                                                                                                                                                                                          19-AUG-1998;
                                                                                                                                                                                                                                                           19-SEP-1997;
                                              Unidentified
                                                                                                                                                              US6294366-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                             25-SEP-2001
                                                                                                                                                                                                                                                                                                                                           Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel B1, Cel B1, Cel 1/2, Cel 1/2/3, Cel 6 or Cel 83/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K635 to N1426 in the sequence shown in AAY13492; Cel B1 extends from amino acid K35 to Cel B1/2 extends from mino acid K35 to G812, Cel E6 extends from amino acid CH124, Cel B6 extends from amino acid CH124, region extends from amino acid CH124, region extends from amino acid CH123, the new sequence shown in AAY13493; Cel E3/B5 is shown in AAY13494. The new
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  enzymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including oction-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   enzymes show reduced redeposition of dye compared to using non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim.
                                                                                                                                                                                                                                                                     Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel B1; Cel B1/2; Cel 1/2/3; Cel 6; Cel B3/B5; laundry detergent; stain-removing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 6415 BP; 2067 A; 1082 C; 1689 G; 1541 T; 0 U; 36 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Farrington GK,
                                                                                                                                                                                                                                   DNA sequence encoding truncated cellulases.
                                                                                                                                                                                                                                                                                                             cotton-containing fabric; stonewashing; ss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Daniels RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Page 24-25; 65pp; English.
                      238 AGTGGAGCGGATTATTACCTGGAG 261
88
81 ProThrAlaAspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         runcated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98EP-00810919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97US-00932571
                                                                                                                  AAX55662 standard; DNA; 6415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  , Bergquist PL, williams DP;
                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-315403/27.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-PSDB; AAY13493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                   Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                          30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                              09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     runcated
                                                                                                                                                    AAX55662;
                                                                                                  AAX55662
```

6415 38 14 32 1

Length:
Matches:
Conservative:
Mismatches:
Indels:

Similarity:

Query Match:

Best Local

Percent Similarity:

Pred. No.:

Score:

2.05e-13 194.50 61.18% 44.71% US-09-917-376-4 (1-89) x AAX55662 (1-6415)

```
Morgan H,
                                                                                                                                                                     Query Match:
                                                                                                                               ..
No
                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                            AAX55661
                                                                                                                                                                                                                                                                                                                                                                                                                                                88888888888888
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                  2724 TGGTTTAAGATAGTGAATGGAGGCAGCAGCAGTGTTGATCTTAGCAGGGTTAAGATAAGA 2783
                                                                                                  TACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2840
                     GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAGGCCG 2723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the
                                                                                                                                                                                                                                                                                                  proteinase;
    23
                                          43
                                                                                63
                                                                                                                      83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
GlyValLy8ValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLy8Pro
                                                                         TyrTrpPheThrArgAspGlyGlyGerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                          24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                     MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                               Active cellulase protein; alkalinophilic; textile processing; protei detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; celE gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gibbs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ď
                                                                                                                                                                                                                                                                                                                                                                                                                    encoding E1/2 protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                /product= "DNA encoding El protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Daniels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ъ,
                                                                                                                                                                                                                                                                                                                                                                                      "CelE protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bergquist
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Col 37-44; 61pp; English.
                                                                                                                                                                                                                                                                              Active cellulase protein, celE gene
                                                                                                                                                                                                                                                                                                                                                          .ocation/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                      路
                                                                                                                                                                                                                                                                                                                                                                                                                   "DNA
                                                                                                                                                                       ||| ||||||:::
GATTATTACCTGGAG 2915
                                                                                                                                                                                                                     AAD26526 standard; DNA; 6416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-00136574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97US-00932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Anderson P,
                                                                                                                                                            AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                    534. .5889
/*tag= a
/product= "
                                                                                                                                                                                                                                                                                                                                                                                                                           748. .2076
                                                                                                                                                                                                                                                                                                                                                                                                 48. .2538
                                                                                                                                                                                                                                                                                                                                                                                                          /*tag= c
/product= '
                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                        *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            containing fabrics, is elevated temperatures
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2002-081780/11.
P-PSDB; AAE16324.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                           26-MAR-2002
                                                                                                                                                                                                                                                                                                                                       Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US6294366-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Farrington
                                                                                                  2784
                                                                                                                                                                              2901
                       2664
                                                                                44
                                                                                                                       64
                                                                                                                                                            84
                                                                                                                                                                                                                                        AAD26526
                                                                                                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                          Key
                                                                                                                                       셤
                                                                                                                                                                             셤
                                                                                                                                                                                                                       d
                                                         g
                                                                                                 g
                                         8
                                                                              8
                                                                                                                      ò
                                                                                                                                                           ò
```

```
celBS amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing whits or for imparting stronwashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2655 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAGGCCG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2725 TGGTTTAAGATAGTGAATGGAGGCAGCAGTGTTGATCTTAGCAGGGTTAAGATAAGA 2784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2785 TACTGGTACACAGTGGAGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E1/2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gibbs MD;
                                                                                                                                                                                                                                                                          Sequence 6416 BP; 2068 A; 1082 C; 1689 G; 1541 T; 0 U; 36 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel B1; Cel Cel 1/2/3; Cel 6; Cel B3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NNA sequence encoding truncated cellulases Cel B4/5 and Cel B5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Farrington GK,
                                                                                                                                                                                                                                                                                                                                                       6416
38
14
32
                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Daniels RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             x AAD26526 (1-6416)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAX55661 standard; DNA; 11706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2902 GATTATTACCTGGAG 2916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98EP-00810919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     97US-00932571
                                                                                                                                                                                                                                                                                                                                                    2.05e-13
194.50
61.18%
44.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΡΓ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84 AspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anderson P, Bergquist
Morgan H, Williams DP;
                                                                                                                                                                                                                           sequence is celE gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1999-315403/27.
P-PSDB; AAY13492.
                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-4 (1-89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unidentified
                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3P921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-JUN-1999
```

Location/Qualifiers

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4, Cel E1/2, Cel I1/2/3, Cel 6 or Cel E3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A101 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A102 to D481, Cel E1/2 extends from mino acid A103 to D481, Cel E1/2 extends from x39 to G812, Cel E1/2 and the stability region extends from amino acid V123 to K1751 and the stability region extends from amino acid A103 to G635, In the sequence shown in AAV13493; Cel E3/B5 is shown in AAV13493; Cel E3/B5 is shown in AAV13491. The new enzymes are useful in laundry detergent compositions to prevent or remove
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nou-
                              New truncated cellulase proteins, useful in detergents and for producing stonewashed' denim.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using r
                                                                                                                     Disclosure; Page 20-23; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            truncated cellulase compositions
```

Sequence 11706 BP; 3828 A; 1994 C; 2994 G; 2890 T; 0 U; 0 Other;

```
4098 TGGTTTAAGATGAATGGAGGCAGCAGCAGTGTTGATCTTAGCAGGTTAAGATAAGA 4157
                                                                                                                                                                                                                                                                        |||||||::|||| |||||||| TACTGGTACACAGAGTGCGGTA---TGTGACTGGGCACAG 4214
                                                                                                                                                                     GGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCGAGCACAAGTTCTATAAGGCCG 4097
                                                                                                                                                                                                                                                                                                                                             GlyValLygValGlnTyrLygAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
                                                                                                                                                                                                    43
                                                                                                                                                                                                                                                          TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                  MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValThrValThrValHrg
              111706
37
16
31
1
              Length:
Matches:
Conservative:
Mismatches:
                                                                       Indels:
                                                                                    Gaps:
                                                                                                               US-09-917-376-4 (1-89) x AAX55661 (1-11706)
             7.82e-13
192.50
62.35$
43.53$
                                                                                                                                                                                                                                                                                                                                                                         AspThrTyrLeuGln 88
                                         Percent Similarity:
Best Local Similarity:
Alignment Scores:
                                                                                                                                                                                                                                                                                                                  64
                                                                                                                                                                                                                                                                                                                                                                          84
                                                                                                                                                                                                                                                                                      4158
                                                                                                                                                                                                                                                           44
                                                                       Query Match:
DB:
                               Score:
                                                                                                                                           8
                                                                                                                                                                    g
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                        ⋧
                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
Active cellulase protein; alkalinophilic; textile processing; proteinase; detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; celB gene; ds.
                                                                                                                                           Active cellulase protein, celB gene.
                                                               BP.
                                                               AAD26525 standard; DNA; 11707
|||| |||||||:::
GATTATTACTTGGAG 4289
                                                                                                                  26-MAR-2002 (first entry)
           4275
                                                                                         AAD26525;
                                     RESULT 12
                                                    AAD26525
         임
```

Unidentified

```
The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the CelB5 amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including octton-containing fabrics, as detergent additives: The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from octton-containing knits or for impatting stonewashed appearance to cotton-containing whits or for impatting are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present sequence is celB gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1158 TACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 4214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4038 GGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCGAGCACAAGTTCTATAAGGCCG 4097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                      or commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                   New cellulase active protein, useful in textile processing or comm detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11707 BP; 3827 A; 1995 C; 2994 G; 2891 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                            Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11707
37
16
31
1
                                                                                                                                                                                                                                                                                                                         Farrington GK, Anderson P, Bergquist P, Daniels R,
Morgan H, Williams DP;
                                                                           /*tag= b
/product= "DNA encoding B4/5 protein"
                                                                                                                                     /product= "DNA encoding B/5 protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                         /product= "CelB protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-4 (1-89) x AAD26525 (1-11707)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Col 27-38; 61pp; English.
                                                                                                   9255. 10526
/*tag= c
                                                                                                                                                                                                                                  98US-00136574.
                                                                                                                                                                                                                                                                                              (CLRN ) CLARIANT FINANCE BVI
                                                                                                                                                                                                                                                               97US-00932571.
              .10535
                                                           8601. .10532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7.83e-13
192.50
62.35%
43.53%
                                /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   elevated temperatures.
                                                                                                                                                                                                                                                                                                                                                                         WPI; 2002-081780/11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                         P-PSDB; AAE16323
                                                                                                                                                                                                                                  19-AUG-1998;
                                                                                                                                                                                                                                                               19-SEP-1997;
                                                             misc_feature
                                                                                                          misc feature
                                                                                                                                                                     JS6294366-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                    25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
 Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
```

63

83

```
|||||||::: ||||||| TACTGGTATAAACGAAAAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 1164
                                                                                                                                               44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                               64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                  24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                               CMCase gene with mutated signal sequence to enhance hydrophobicity.
                                                                                                                                                                                                                                                        AAL41025 standard; DNA; 1488 BP
                                                                                                                                                                                                        1225 GATACCTATCTGGAA 1239
                                                                                                                                                                                 84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                       Protein surface-displayed;
high-throughput screening;
                                                                                                                                                                                                                                                                                                                                                                                           Unidentified
                                                                                                                                                                                                                                                                                                        11-OCT-2002
                                                                                                                                                                                                                                                                      8
                                                        원
                                                                                  ò
                                                                                                       g
                                                                                                                                  ò
                                                                                                                                                      셤
                                                                                                                                                                                ò
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to methods for preparing a protein surfacedisplayed on a genetic carrier, improving a protein or isolating a substance in a mixture comprising transforming a host cell harbouring a genetic carrier consisting of a spore or virus with the vector library. The method is useful for surface display of proteins, particularly for displaying a protein on the surface of e.g. spores, for improving a protein, or for isolating a substance. These methods are useful in obtaining monoclonal variants from a large library, or in high-throughput screening of antibodies for use in therapy (e.g. as vaccines), diagnosis or analysis. This polynucleotide sequence represents the CMCase gene with
Methods for surface display of proteins, e.g. displaying a protein on the surface of spores, improving a protein or isolating a substance, by transforming a host cell harboring a genetic carrier with spore or virus
                                                                                                                                                                                            Protein surface-displayed; genetic carrier; spore; monoclonal variant; high-throughput screening; antibody; vaccine; enzyme; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               additional His encoding sequences relating to the invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1434 BP; 493 A; 293 C; 341 G; 307 T; 0 U; 0 Other;
                                                                                                                                                                      CMCase gene with additional His encoding sequences.
                                                                                                                                                                                                                                                                                             "CMCase protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example 5; Page 109-113; 118pp; English.
                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                               AAL41028 standard; DNA; 1434 BP
                                    ||| |||||:::
4275 GATTATTACTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                     15-JAN-2002; 2002WO-KR000059
                                                                                                                                                                                                                                                                                                                                                                                             15-JAN-2001; 2001KR-00002156
                       AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                           Choi SK, Jung HC;
                                                                                                                                                                                                                                                                                 /*tag= a
/product= '
                                                                                                                                              11-OCT-2002 (first entry)
                                                                                                                                                                                                                                                                      L. .1434
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                     (GENO-) GENOFOCUS CO LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             with the vector library
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-590661/63.
P-PSDB; AAO22444.
                                                                                                                                                                                                                                                                                                                     WO200255561-A1
                                                                                                                                                                                                                                   Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                             18-JUL-2002
                          84
                                                                                                                       AAL41028
                                                                                                                                                                                                                                                                                                                                                                                                                                           Pan JG,
                                                                       RESULT 13
                                                                                                                                                                                                                                                           Key
                                                                                     AAL41028
                                                                                               ò
```

genetic carrier; spore; monoclonal variant; antibody; vaccine; enzyme; gene; ds.

(first entry)

Location/Qualifiers

```
displayed on a genetic carrier, improving a protein or isolating a substance in a mixture comprising transforming a host cell harbouring a genetic carrier consisting of a spore or virus with the vector library. The method is useful for surface display of proteins, particularly for displaying a protein on the surface of e.g. spores, for improving a protein, or for isolating a substance. These methods are useful in obtaining monoclonal variants from a large library, or in high-throughput sorreming of antibodies for use in therapy (e.g. as vaccines), diagnosis or analysis. This polymucleotide sequence represents the CMCase gene with a mutated signal sequence to enhance its hydrophobicity relating to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Methods for surface display of proteins, e.g. displaying a protein on the surface of spores, improving a protein or isolating a substance, by transforming a host cell harboring a genetic carrier with spore or virus with the vector library.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to methods for preparing a protein surface-
                                                                                       /product= "CMCase protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 4; Page 98-103; 118pp; English.
                                                                                                                                                                                                                                                                                                                                                15-JAN-2002; 2002WO-KR000059
                                                                                                                                                                                                                                                                                                                                                                                                                                   15-JAN-2001; 2001KR-00002156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Choi SK, Jung HC;
1. .1491
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GENO-) GENOFOCUS CO LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2002-590661/63.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P-PSDB; AA022443
                                                                                                                                                                          WO200255561-A1
                                                                                                                                                                                                                                                             18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pan JG,
```

4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23

1434 34 18 31 1

Length: Matches: Conservative: Mismatches: Indels:

1.94e-13 188.00 61.18% 40.00%

Best Local Similarity:

Query Match: DB:

Percent Similarity:

Gaps:

US-09-917-376-4 (1-89) x AAL41028 (1-1434)

ò

us-09-917-376-4.p2n.rng

```
Alignment Scores:
                                                                                                                                                   microarray
                                                                                                                                                                                                                                                                                                                                               invention
                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                            Pan J,
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                          TACTGGTATAACGCGAAAAACAAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 1221
                                                                                                                                                                                                                                 :::|||||||||||::
CTTGGATGCGGCAATGTGACATACAAGTTTGTGATGCATAAACCAAAGCAAGGTGCA 1281
                                                                                                                         43
                                                                                                                                                                                   TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                      64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                              GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
                                                                                                                                                 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                               Spore coat protein; spore surface; delivery system; bioconversion; recombinant live vaccine; resistence property; protein microarray; whole cell absorber; secretion signal; target signal; cell membrane; surface display motif; cotG; CMCase; carboxymethyl cellulase; gene;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /product= "Bacillus subtilis spore coat protein c
carboxymethyl cellulase (CMCase) fusion protein"
/note= "This sequence lacks a start codon"
                                                                                                                                                                                                                                                                                                                                                              Bacillus subtilis spore coat protein cotG-CMCase fusion gene.
        357 G; 333 T; 0 U; 0 Other;
                                 1488
34
18
31
1
                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /*tag= e
/note= "CMCase structural gene"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /*tag= c
/note= "CotG structural gene'
                                 Length:
Matches:
                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /*tag= a
/note= "CotG promoter"
                                                                                             US-09-917-376-4 (1-89) x AAL41025 (1-1488)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1046. .1084
/*tag= d
/note= "Linker"
         ΰ
                                                                                                                                                                                                                                                                        1282 GATACCTATCTGGAA 1296
         501 A; 297
                                                                                                                                                                                                                                                                                                            ABK53202 standard; DNA; 2510
                                2.04e-13
188.00
61.18%
40.00%
                                                                                                                                                                                                                                                        AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        085. .2491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61. .1045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61. .2494
                                                                                                                                                                                                                                                                                                                                              12-AUG-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         *tag=
                                                                                                                                                                                                                                                                                                                                                                                                                          Bacillus subtilis.
                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
        Sequence 1488 BP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200246388-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                    1168
                                                                                                                                                                                                                                      1222
                                                                                                               4
                                                                                                                                                 24
                                                                                                                                                                                    44
                                                                                                                                                                                                                                                        84
                                                                                                                                                                                                                                                                                                                                                                                                                                             Chimeric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      promoter
                                                                                                                                                                                                                                                                                         RESULT 15
ABK53202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                             ઠે
                                                                                                                              ద
                                                                                                                                                ò
                                                                                                                                                                 ద
                                                                                                                                                                                  ò
                                                                                                                                                                                                   유
                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                       셤
```

```
The present invention relates to a new method for displaying a protein of interest on a spore surface. The method of the invention involves preparing a vector comprising a gene construct containing a gene encoding a spore coat protein and a gene encoding a protein of interest, transforming a host cell with the vector, displaying the protein of interest on a surface of a spore of the host cell, and recovering the spore. The method is useful for producing a delivery system for recombinant live vaccine. The method is also useful for improving a recombinant live vaccine. The method is also useful for improving a notein of interest using a resistance property of spore, in bioconversion reactions, for preparing protein microarray, for producing an antibody to antigen in vertebrates, and for preparing a whole cell absorber. The spore coat proteins used circumvent a necessity for passage across cell membrane, so that they do not need secretion signal and target signal which are prerequisites of surface display motif, therefore no surface display of protein, such as beta-galactosidase, in orderly fashion, which is difficult to pass across cell membrane. The protein cotd-carboxymethyl cellulase (CMCase) fusion protein of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2171 TACTGGTATAACGCGAAAAAAAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83
                                                                                                                                                                                                                                                                                                                                                         Displaying a protein of interest on spore surface, by transforming a host cell with a vector comprising a gene encoding a spore coat protein and a gene encoding a protein of interest, useful for preparing protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2510 BP; 869 A; 528 C; 521 G; 592 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2510
34
18
31
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-4 (1-89) x ABK53202 (1-2510)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 4; Page 145; 158pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-DEC-2001; 2001WO-KR002124
                                                             08-DEC-2000; 2000KR-00074835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.97e-13
188.00
61.18%
40.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84 AspThrTyrLeuGln 88
                                                                                                                             (GENO-) GENOFOCUS CO LTD
                                                                                                                                                                                                                                                          WPI; 2002-463830/49.
P-PSDB; AAU98063.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
                                                                                                                                                                                             Choi S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
```

Search completed: May 12, 2004, 05:39:24 Job time : 134.455 sec8

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
OM protein - nucleic search, using frame plus p2n model
```

May 12, 2004, 04:33:52 ; Search time 835.94 Seconds

Run on:

	<pre>(without alignments) 3179.336 Million cell updates/sec</pre>
Title: Perfect score: Sequence:	US-09-917-376-4 470 1 VSGGVKVQYKNNDSAPGDNQRASFGSVNPATPTADTYLQX 89
Scoring table:	BLOSUM62 Xgapop 10.0 , Xgapext 0.5 Ygapop 10.0 , Ygapext 0.5 Fgapop 6.0 , Fgapext 7.0 Delop 6.0 , Delext 7.0
Searched:	27513289 segs, 14931090276 residues

55026578 Total number of hits satisfying chosen parameters:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2 1/USFTO spool/US09917376/runat 11052004 114535 28249/app query.fasta_1.1429
-Q=/cgn2 1/USFTO spool/US09917376/runat 11052004 114535 28249/app query.fasta_1.1429
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCL=0 -LOOFEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blooum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=560 -MINLEN=0 -MAXLEN=200000000
-USER=US09917376 @GGN 1 1.5334 @runat 11052041 114535 28249 -NCFU=6 -ICFU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO MAT -LARGEQUERY -NEG SCORES=0 -MAIT -DSPBLOCK=100 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

em_gss_hum:* em_gss_inv:* em_gss_pln:* em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gssl:* em gss vrt:* em gss fun:* em_gss_fun: em_gss_mam: em_estba:*
em_esthum:*
em_esthum:*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estro::*
em_estr em_estfun:* em_estom:* em_gss_mus: EST: * Database :

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

o.	cription	30372 POL1 65	17524	3085 UCR	722991	3716	2891	6525	282(649	392	32	45	80	741	74	177	ò	491464	CA595709 Wpalc.pk	7010	777	400400	DF8/39/08 KC3-ELUIS CR3943980 IICRCS01 0	3084	1904	850	3232	38 C64738 Yuj	22108 laf13d	9426 Danio	9803 CH230	7853 PUFKJ8	7856 PUFKJ8	16375 dad35b	9805 AL799805	5312	43 ALB98643	2471	54312 EST-77-	2576	5360	5493 AL846493	BM609423 170006871
	Ω	: 8	752	308	299	7.1	289	552	282	764	939	532	345	280	74	7.4	312	ָרָרָ מַרְי	146	270	# (7 0	200	6F8/5908 CB262980	308	190	5585	823	38	2	945	980	85	785	81637	9980	9631	9864	4247	75431	5257	2536	93	5094
	DB	14		14	13		10	14	10	14	12	78	12	58	53	23	10	14	10	14	2 .	1.4		10		12	13	14	13	14	53	78	29	29	12	σ	10	σ	28	14	σ	29	6	12
	Length		754	768	906	376	909	641	657	919	691	724	814	905	σ.	1001	9	1058 1058	465	613	000	302	500	107	819	0	1237	357	360	593	116	861	894	924	380	450	533	544	563	267	579	601	640	648
ď	ery	17.2	9	v	ė	9	è.	è.	ė.	è.	ė.	ġ.	ė.	ė	٠.	٠	٠ ف	۰	'n	٠ د		Λu	n	15.6		S	S.		'n.	S.	S.	ŝ	ů.	ů.	ů.	ů.	'n	S	S	15.4	S	15.4	ß	15.4
	cor	81	6	78.5	7	16	16	92	94	92	92	16	92	16	16	7		'n		74.5	. [4, 5	٠,	7.5.5		, m	Ε.	73	73	73	73	73	73	7	ď	72.5	'n.	'n.	ς.	ς.	ς.	ζ.	ς.	ς.
	Ξ.ο	-	~	m	4	Ŋ	9	7	œ	0	10	c 11	12	-	c 14	15	٦,	-		c 13	N (77.	v c	57 C	25.	ı ca	c 27	28	29	30	31	32	c 33	34	m	c 36	37	m	39	40	c 41	42	4	c 44

ALIGNMENTS

		CF480372 584 bp mRNA linear EST 08-SEP-2003			CF480372.1 GI:34509241	EST.	Sorghum bicolor (sorghum)		Bukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD	clade; Panicoideae; Andropogoneae; Sorghum.
RESULT 1	CF480372	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM			

```
BQ047524 10-MAR-2003
EST596642 P. infestans-challenged potato leaf, incompatible
reaction Solanum tuberosum cDNA clone BPLI17P18 5' end, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

( Dases 1 to 754)

Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukoianov, A., Rangel, P., Haberlach, G.T., Karamycheva, S.-A., Tsai, J., Chiemingo, A., Generation of ESTs from Potato Leaves Challenged with Phytophthora infestens, incompatible Interaction (2002)

Contact: Robin Buell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: UC Berkeley, PGEC; sequencing: The Institute for Genomic Research. Whole plants were challenged with 450,000 sporangia/ml P. infestans isolate US-1 (US940501) in Biotron (Madison, Wisconsin). Leaf tissue was collected at 1, 2, 5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
      213 GCTTGTGTGAAGTGTGGTGACGTTGGTTCTTTGGAGACGGTGTCTTGCTCTGTTGGTCCG 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7 ValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGln 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato. "
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /db xref="taxon:4113"
/clone="BPL:117P18"
/tissue type="leaf"
/dev_stage="6 week old"
/lab host="SOLR"
/clone_lib="P. infestans-challenged potato leaf,
incompatible reaction"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                754
29
114
33
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           nttp://genome.arizona.edu/orders/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89) x BQ047524 (1-754)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seq primer: T3.
Location/Qualifiers
1. .754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Solanum tuberosum (potato)
                                                                                                        273 AGCTTGCAGTCTGGTGAT 290
                                                                    84
                                                                                                                                                                                                                                                                                                                                                                                   BQ047524
BQ047524.1 GI:19821510
                                                                 79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27.9
79.50
38.05%
25.66%
16.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Solanum tuberosum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                            sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local S.
Query Match:
                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                          LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
COMMENT
                                                                                                                                                                                            RESULT 2
BQ047524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                              g
                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /mol_type="mRNA"
/colltiva=="mRNA"
/colltiva=="mRNA"
/colltiva=="mRNA"
/colltiva=="mRNA"
/do_type="mRNA"
/clone="mpOil 65 H07 A002"
/lab_host="npH10B-T1" phage-resistant E. coli"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen"
/clone_lib="mpOilen
                                                                                                                                                                                                                                                                                                                                                                                                                                       Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & W University,
Science; plant material and RNA prepared at Texas A & W University,
Sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude.polyA.
Seq primer: Sug3-14 (TAGTCTAGCGGCGCGACC)
POUNA=Yes.
                                   Liang, C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              105 ------CGCGCCGGACAAACCCTGAAACGATCGATGGTGCAAGGAAA 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCCTGTGCTGGCCCGGGC 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
         cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J., Lucas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and Pratt, L.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                           EST database from Sorghum: pollen Unpublished (2003)
Other ESTs: POL1 65 H07 A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             584
24
15
37
10
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-4 (1-89) x CF480372 (1-584)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12.7
81.00
45.35%
27.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
REFERENCE
AUTHORS
                                                                                                                                                         TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                      COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

셤

ò

g

8

g

셤

ò

ò

셤

ò

```
/clone="UVCKCS01" (2012")
// issue_type="Rind containing flavedo and albedo"
// dev stage="Mature fruit"
// lab host="E. coli TUCI21"
// lab host="E. coli TUCI21"
// lab host="E. coli TUCI21"
// clone_lib="Washington Navel orange cold acclimated
flavedo & albedo cDNA library.
// note="Vector: Lambda Uni-ZAP XR, excised phagemid;
// site 1: EcoRI; Site 2: XhoI; Plants were grown in the
field at University of California, Riverside Agricultural
Operations since 1983. The scion was Washington Navel
orange and the rootstock Carizzo Citrange. Tissue from
mature fruit was harvested at mid-day in January 2002
during a cold spell, when pre-dawn temperatures were
approximately -2 to 2 degree C. Approximately 2 cm median
sections of the rind were excised in the field from
several fruits, then wrapped in aluminum foil and frozen
quickly in dry ice. Total RNA was extracted using a phenol
extraction procedure described in J. Japanaese Soc. Hort.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CB293085 USCCS01 05cb12 gl Washington Navel orange cold acclimated flavedo & albedo cDNA library Citrus sinensis cDNA clone UCRCS01_05cb12, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Feuerbacher, O., Kim, H.R.,
Yu, Y.
Genetic Markers for California
                                                                                                                                                                                                                        269 AAGGGGGAGCCTCCGAAGTTTCTTGGTTATGAACATTTCTCCACTCGCAAATGGGGTTCA 328
                                                                                                                                                                                                                                                                                                                                  388
                                                                                                                                                                                                                                                                                                                                                                                                                            Bukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta; Magnoliophyta, eudicotyledons, core eudicots,
rosids, eurosids II; Sapindales, Rutaceae, Citrus.
                                                                                                             268
GTGCCTTATCAAGATCCAGGAAGT----CCGGGTAGTAATCTAATATCTCCTGGTTCAGTA 208
                                                                                                                                                                                                                                                                               54
                                                                                                                                                                                                                                                                                                                                                                                         74
                                                                                                                                                              -----ValAspLeuSerThrValThrValArgTyrTrpPheThr 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
                                                                                   329 ÁGGGTTÖGÖTCAGGATCATTGACACCAAGTGGCTGGGGCTCTAGGCTAGGTTÖTGGAACT
                                                                                                                                                                                                                                                                            48 ArgAspGly-----GlySerSerThr
                                                                                                                                                                                                                                                                                                                                                                                         55 LeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         75 SerValAsnProAlaThrProThrAlaAspThrTyrLeu 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /mol_type="mRNA"
/cultivar="Washington navel"
/db_xref="taxon:2711"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 768)
Close, T.J., Collura, K., Fenton, R.D.,
Kudrna, D., Wanamaker, S., Wing, R. and
Development of EST Resources and New (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       organism="Citrus sinensis"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tel: 9097873318
Fax: 909787437
Email: timothy.close@ucr.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                      27 LeuValAsnThrGlySerSer-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CB293085.1 GI:28618542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unpublished (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Citrus sinensis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seq primer: T3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Citrus sinensis
                                                                                                                                                                          ------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Citrus
                                                                                                                                                                    35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CB293085
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
```

```
Sci. (1996) 64:809-814. Poly(A) RNA was purified, a CDNA library was made, and 1 million primary lambda CDNA clones were in vivo excised to give a population of pBluescript SK(-) phagemids. All steps to this point were performed in the TJ Close lab at the University of California, Riverside (Fenton). Phagemids were plated, plasmid DNA purified, CDNA clones archived, and DNA sequences determined bi-directionally using an AB13530 at the Arizona Genomics Institute, University of Arizona (Collura, Feuerbacher, Kim, Kudrna, Wing, Yu). Chromatogram files were transmitted to UC Riverside (by Yu), then processed at UC Riverside (by Wanamaker) using the HarvEST pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DQ/12391
AGENCOURT 8103733 Lupski sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6191052 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              663 AATAATAGCGACGTTGTTACTAGTGGTGAGCGCTAGCGTTGGTGCAAGTGCGGTGACG--- 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                612 GGCGCCATCAAC-----AAGAACAACAACAGCAGCAGCAATAATAATAAAAATAAT 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 906)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23 ProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrVal 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I. M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information found through the I.W.A.G.E. Consortium/LLNL at:
http://mage.llhi.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      768
20
10
15
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43 ArgTyrTrpPheThrArgAspGlyGlySerSerThr 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                High quality sequence stop: 512.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x CB293085 (1-768)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BQ722991.1 GI:21861888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37.9
78.50
57.69%
38.46%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Si
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BQ722991
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠ
```

```
Sorghum bicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BE362891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BE362891
                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                         /.clone_lib=Tuppski sympathetic trunk"
//clone_lib=Tuppski sympathetic trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_l:
Not1; Site_2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5.-TCGACCCAGGGGTCGG-3 and
5.-GACTAGTTCTAGATCGGGGGGGCGCCCT(15)-3'. Size selected >
1.kb for average insert length 1.9 kb. This is a primary library, non-amplified Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           zk61f10.s1 Soares pregnant_uterus NbHPU Homo sapiens CDNA clone IMAGE:487339 3' similar to_gb:L10240 BASIGIN PRECURSOR (HUMAN);,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryotzi, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi, Mammalia; Butheria, Primates; Catarrhini, Hominidae, Homo.

1 (bases, 1 to 37).

Hillier, L. Lennon, G. Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hulman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Moore, B., Mostes, M.B., Tan, F., Thierry-Meg, J., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Meg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             548 CTGCGCTATACCAATAAAATCTCCAGCGAGGCCACCGTGAGGTA-----ATGAAG 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7 ValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGln 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 LeuvalAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrpPhe 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TCCCGG---GGCGCCAGCTCCTACACTGCATCTGGGGCAGTGGTGGT 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 47 ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly 65
                                                                                                                                                                                                                                                                                                                                                                                  906
116
113
26
2
                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
Conservative:
Mismatches:
Indels:
                                     /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
clone="IMAGE:6191052"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            x BQ722991 (1-906)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA043716.1 GI:1521725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                  73.1
77.00
49.15$
27.12$
16.38$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            and Marra, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97044478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA043716
                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3889549
                                                                                                                                                                                                                                                                                                                                                                 Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                  No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                 Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
AA043716
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                    Pred.
```

ò g ò 셤 ò 원

```
/ Gex="female" | / Gex="female" | / Gex="female" | / Gev stage="adult" | / Jab host="Bulton" | / Jab host="Bulton" | / Jab host="Bulton" | / Jab host="Bulton" | / Jab host="Dispersion | / Jab host="Dispersion | / Jab host="Dispersion | / Jab host="Dispersion | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Jab host | / Ja
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BE362891 606 bp mRNA linear EST 17-DEC-2001
DG1 90 C11.gl A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           178 CCCCTCCAGAGCTGCCCCCGGATGCACAAGGAGGGGGGCTTGGGAGTGACTTGAGGC 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukāryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlyAspAsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAsp 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----LeuValTyrAsn 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37 LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThr----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 (bases 1 to 606)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pratt, L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
On July 20, 2000 this sequence version replaced gi:9304448.
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TGTGACGGGGTCGCCCTCGGGTGTGGGCAAGTGAAGTCCTCTGTGGGC 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               376
20
13
23
20
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
Insert Length: 1400 Std Error: 0.00 Seq primer: -40M13 fwd. from Amersham High quality sequence stop: 212. Location/Qualifiers
                                                                                                                                                                                                                                               /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                  /mol_type="mRNA"
/db_xref="GDB:3761053"
/db_xref="taxon:9606"
/clone="IMAGE:487339"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JS-09-917-376-4 (1-89) x AA043716 (1-376)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BE362891.2 GI:15724434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28.5
76.00
43.42%
26.32%
16.17%
```

```
вызогиго
DG1 89 C02.g2_A002 Dark Grown 1 (DG1) Sorghum bicolor CDNA, mRNA
sequence.
                                                                                                                                                                   Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science, plant material and RNA prepared at Texas A & M University,
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia, Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         149 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCCTGTGCTGGCCCGGGC 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     209 ------CGCGCCGGAACCCTGAACTTCCAAACGATGGTGCAAGGAAA 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              317 GCTTGTGTGAAGTGTGGTGGCGTTCTTTGGAGGAGGGTGTCTTGCTCTGTTGGTCCG 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /clone_lib="Callus_culture/cell suspension"
/note="Vector: pWE18S-FL3; Site 1: Xho1; Site_2: Xho1; Th.
library was prepared from a mixture of poly4+RNA from
callus culture tissue and cells in suspension culture.
Double-stranded cDNA was cloned unidirectionally into
different DraII sites of the pWE18S-FL3 vector (5-prime
DraII site is CACGTGTGTG, 3-prime DraIII site is
CACCATGTGT). Xho1 excises the CDNA insert."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
Other ESTs: CCC1_46_E07.gl_A007
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="taxon:4558"
/clone="CCC1_46_E07_A007"
/lab_host="DH10B-T1_phage-resistant_E._coli"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                               Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
POLYA=Yes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="RTx430"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-4 (1-89) x CD226529 (1-641)
                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ::: :: :: |||
377 AGCTTGCAGTCTGGTGAT 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BE362820
BE362820.1 GI:9304377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59.9
76.00
44.19%
26.74%
16.17%
                                                                                                                            Tel: 706 542 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                          .641
                                                                                                                                                                                                                                                                                                                                                                  exclude polyA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BE362820
                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CD226529

CCC1_46_E07.bl_A007 Callus culture/cell suspension Sorghum bicolor CDNA clone CCC1_46_E07_A007 3', mRNA sequence.
                                                                                            Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTMix or 7 sequencing primer, are presented as the reverse complement. Seq primer: PolyTMix High quality sequence start: 15
High quality sequence stop: 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 268 -------CGCGCCGGGACAGAACCCTGAAACTTCCAAACGATCGCTGCAAGGAAA 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sorghum bicolor
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                       /clone_lib="coart Grown 1 (DG1)"
/clone_lib="coart Grown 1 (DG1)"
/note="Corgan: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_l: Xhol; Site_2: BcoRl; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCAGCTGTGCTGGCCCGGGC 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Rathore, K., Eastman, A. and Pratt, L.H.
An EST database from Sorghum: callus culture and cell suspension Unpublished (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
       Rm. 2502, Athens, GA 30602-7271, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        606
23
115
110
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                       'organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                             /mol_type="mRNA"
/db_xref="taxon:4558"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          x BE362891 (1-606)
                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sorghum bicolor (sorghum)
  Plant Sciences Building,
Tel: 706 542 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36 AGCTTGCAGTCTGGTGAT 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GI:30969963
                                                                              Email: mmpratt@uga.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (bases 1 to 641)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55.4
76.00
44.19%
26.74%
16.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CD226529.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-4 (1-89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
                                                                                                                                                                                                                                                                          POLYA=No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CD226529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CD226529
                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
```

ò

ð 유 셤 ò a

ઠ

g

ò

ORIGIN

641 23 15 38 10

43

78

```
/organism="Sorghum bicolor"
/wol type="mRNA"
/wol type="mRNA"
/wol type="mRNA"
/db xref="taxon:4558"
/db host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl108-T1 phage-resistant E. coli"
/lab host="mBl11 in Bl108-T1 phage-resistant E. coli"
/lab host="mBl11 site is CACCATGTG].
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host="mBl11"
/lab host
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              214 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGAGGAGCCTGTGCCTGGCCCGGGC 273
                                                                                      Sorghum bicolor
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicolees; Andropogoneae; Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       274 -------CGCGCCGGGACAGAACCCTGAAACTTCCAAACGATCGGTGCAAGGAAA 321
                                                                                                                                                                                                               1 (bases 1 to 676)
Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J., Lucas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 583 0210
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 676
23
15
10
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Seq primer: Sug3-14 (TAGTCTAGCGGCCGCGACC)
POLYA=Yes.
                                                                                                                                                                                                                                                                                                                                         Pratt, L.H.
EST database from Sorghum: pollen
Unpublished (2003)
Other ESTs: Poll 45_C08.91_A002
Contact: Cordonnier-Pratt MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-4 (1-89) x CF487649 (1-676)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                               Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Email: mmpratt@uga.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64.6
76.00
44.19%
26.74%
16.17%
   CF487649.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                      REFERENCE
AUTHORS
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST 08-SEP-2003
                                                         Sorghum bicolor
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /mol_type="mRNA"
/db_xref="taxon:4558"
/db xref="taxon:4558"
/clone lib="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site 1: Xhol; Site 2: ECORI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              204 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGAGGAGGAGCCTGTGCTGGCCCGGGC 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 264 ------CGCGCCGGGACCGTGAACTTCCAAACGATGGTGCAGGAA 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44 ---TyrIrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                                                                                                                                                               Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Fax: 706 583 0210
                                                                                                                                                                                                                   Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 676 bp mRNA linear EMPOLI 45_C08_b1_A002 Pollen Sorghum bicolor cDNA clone POLI 45_C08_A002 3', mRNA sequence.
                                                                                                                                                                                                                                                                                  An EST database from Sorghum: dark-grown seedlings
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     657
23
15
38
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Seg primer: PolyTMix
High quality sequence start: 89
High quality sequence stop: 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-4 (1-89) x BE362820 (1-657)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                               Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGCTTGCAGTCTGGTGAT 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlaThrProThrAlaAsp 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: mmpratt@uga.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
76.00
44.19%
26.74%
16.17%
                                                                                                                                                                                            (bases 1 to 657)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLYA=No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 432
                         SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                         TITLE
JOURNAL
COMMENT
                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
CF487649
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
```

셤 ò 셤 à ద ò 셤 ò 24

g ò g ð

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; Eukaryota; Varidiplantae; Streptophyta; Eukaryota; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (Dases Panicoideae; Andropogoneae; Zea.

1 (Dases 1 to 724)

2 Ratzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8884

Fax: 516 367 8884

Fax: 516 367 8884

Cold Springer Column: 08

Seq primer: -21M13UnivRev

Class: Shotgun

Tel: 52 Annenber Column: 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BZ616323 16-JAN-2003 ig58e08.gl WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone ig58e08 5', genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The vector was digested with XbaI and one nucleotide was added by fill in in the recessive all end one nucleotide was was nabulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in Ml3mpl9, .b/g reads in pUC19). The same ligation was transformed into DBSa."
                                                                                                                                                                                                                      325 -------CGCGCCGGGACAGAACCCTGAAACTTCCAAACGATCGGTGCAAGGAAA 372
                                                                                             62 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                          44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     High quality sequence stop: 728.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:4577"
/clone="ig58e08"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-4 (1-89) x BZ616323 (1-724)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             lab_host="DH5a"
                                                                                                                                                                                                                                                                                                                                           493 AGCTTGCAGTCTGGTGAT 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BZ616323
BZ616323.1 GI:27766878
                                                                                                                                                                                                                                                                                                      79 AlaThrProThrAlaAsp 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71
76.00
36.28%
30.09%
16.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
BZ616323/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                     셤
                                                                                                                                                                                                                                     g
                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                      ઠે
                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Sorghum bicolor"
/mol type="mRNA"
/mol type="mRNA"
/db xref="taxon:4558"
/clone lib="Embryo (EM1)"
/note="Organ: Embryo germinated for 24 hr; Vector:
/bBluescript II from Lambda Zap II; Site 1: XhoI; Site 2:
EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."
                                                                                                                                                                                                                                                                                                                                                                                                                                                     BG739392 691 bp mRNA linear EST 15-MAY-2001
EM1 82_D05.g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota (Midiplantae) Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD cade; Panicoideae; Andropogoneae; Sorghum.

( Dases 1 to 691)

Reid, S. P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.

The SST database from Sorghum: developing embryos

Umpublished (2000)

Contact: Cordonnier-Pratt MM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                324
                                                                                                                                                                                      322 ccargericioscarcrocaiciosaaciais de de cargeraria de cargera de de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de cargera de carger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24
                                  61
                                                                                                                                                    78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                    AlaalaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   691
23
15
15
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seq primer: PolyTWix
High quality sequence start: 50
High quality sequence stop: 691
POLYA=No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-4 (1-89) x BG739392 (1-691)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sorghum bicolor (sorghum)
Sorghum bicolor
                                                                                                                                                                                                                                                                                                               442 AGCTTGCAGTCTGGTGAT 459
                                                                                                                                                                                                                                                                             84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BG739392.1 GI:14089081
                                                                                                                                                                                                                                                                          79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66.6
76.00
44.19%
26.74%
16.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        is 20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Scores:
                                                                                                                                                    62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25
                                     44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
BG739392
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
```

g 8

79

d ò g ò 엄 ò

```
אנבים אום 105 bp DNA linear GSS 16-DEC-2002 OGALR72TM ZM2 0.7 1.5 KB Zea mays genomic clone ZMMBMa0089L24, genomic survey sequence.
                                                                                                                                                                 117 CTCCGCTACTACAACCAGAGCGCGGGGGGCTCTCACACGATCCAGCGGATGTACGGCTGT 176
                                                                                                                                                                                                                                                                                                                                          177 GACGTGGGGTCGGACGGCCC-----TCCTCCGCGGGTACGAGCAGGTCGCA 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Zea.
                                                                                                                                      22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 (bases 1 to 905)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSS: OGALR72TC
                                                           2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                  60 Asp-TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAl
                                                                                                                                                                                                                      42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeu-----ValTyrAsnCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. :905.

/organism="zea mays"

/mol_type=genomic DNA"

/strain=B73"

/db_xref="taxon:4577"

/clone="xmmna0089124"

/clone="vector: DBCSK-; Site_1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9712 Medical Center Drive, Rockville, MD 20850, Tel: 301-838-5843
Fax: 301-838-0208
Email: TW
Seq primer: TR
Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89) x BZ545807 (1-905)
                    US-09-917-376-4 (1-89) x BI218457 (1-814)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                       225 TACGACGCTGCGATTACA 243
                                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BZ545807.1 GI:27099699
GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Cathy Whitelaw
TIGR
                                                                                                                                                                                                                                                                                                                                                                               79 aThrProThrAlaAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97
76.00
36.28%
30.09%
                                                                                                                                                                             ----- 06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zea mays
Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BZ545807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
BZ545807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                      δ
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BI218457 BI4 bp mRNA linear EST 11-JUL-2001
                                     587 riciocaddicarcaearcccaracaaddaadacaacaddaacaccrocogcagordcocci
                                                                                                     467 ACGAGGAAGTTCGTCAACTTCCCGGAGATCGTCGGAGCGGTCCAGAAGGCTGGGTTCGAG 408
                                                                                                                                                                                                                                                                              407 GTCATCCCGATCGAGCCGCGTAGAGACCTCAGCGTGGAGGACTTCGCCCGGACCGTCGAC 348
                                                                                                                                                                                                                                                                                                                                                             317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus. 1 (bases 1 to 0114) NIH-WGC http://mgc.nci.nih.gov/.
                                                                                                                                                             36
                                                                                                                                                                                                                                           23
                                                                              53
                                                                                                                                                                                                                                                                                                                   53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh 73
  -----AlaPro 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db.cref="taxon:10090"
/db.clone="INAGE:5101069"
/lab_host="DH108 (T1 phage-resistant)"
/clone lib="NGI CGAP_Li9"
/note="Organ: liver; Vector: pCWV-SPORT6; Site_1: NotI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Email: cgapDs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Flate: LLAMILS I row: E column: 14

High quality sequence stop: 792.
                                                                                                                                                                                                                                     37 ------LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSe
                                                                              Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn-----
                                                                                                                                                                                                                                                                                                                                                             347 TCGTGCG----ACGTGCTCATGGGCGCGCACGGGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     814
28
111
23
19
                                                                                                                                                                                                                                                                                                                                                                                                                         316 ------CGGGCCTCACCAACTTCTTCC 293
SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer--
                                                                                                                                                                                                                                                                                                                                                                                                    73 eGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. .814
/organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                          30 Thr-GlySerSerServalAsp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /mol_type="mRNA"
/strain="FVB/N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BI218457.1 GI:14671901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.6
76.00
44.83%
32.18%
16.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished (1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BI21845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EST
                                                                              17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
```

ACCESSION

VERSION

FEATURES

Score:

ORIGIN

RESULT 12 BI218457

g

ò g ò ద us-09-917-376-4.p2n.rst

원 8 요 ò 엄 ò 원

```
CG270748 1001 bp DNA linear GSS 25-AUG-2003
OGWIB33TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0583B18,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

(base, 1 to 1001)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Cineck, R.W., Nunberg, A., Robbins, D. and Lakey, N.
                                           624 GTCATCCCGATCGÁGCGCGTÁGAGACCTCAGCGTGGAGGÁCTTCGCCCGGACCGTCGAC 565
                                                                                                                                   744 elecococadoradoracoracidadecadocacadoracidarecitaroa de consecos 685
                                                                                                                                                                                                                   684 ACGAGGAAGTTCGTCAACTTCCCGGAGATCGTCGGAGCGGTCCAGAAGGCTGGGTTCGAG 625
                                                                                                                                                                                                                                                                                                                                                                               ----- 534
                                                                                              Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn----- 29
                                                                                                                                                                                                                                                          37 ------LeuSerThrValTrValArgTyrTrpPheThrArgAspGlyGlySerSe 53
                                                                                                                                                                                                                                                                                                                                    53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh 73
                  .-----18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /clone="ZMMBMa0583E18"
/clone lib="ZM 0.7 1.5 KB"
/note="Vector: PBCSK-; Site 1: HincII; 0.7-1.5
methylation filtered genomic DNA library"
                                                                                                                                                                            30 Thr-GlySerSerValAsp------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9712 Medical Center Drive, Rockville, MD 20850,
                                                                                                                                                                                                                                                                                                                                                                             564 TCGTGCG-----ACGTGCTCATGGGCGCGCACGGGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                            73 eGlySerValAsnProAlaThrProThrAlaAspThr 85
                  SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                db_xref="taxon:4577"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CG270748
CG270748.1 GI:34182889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Other GSSs: OGWIB33TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1112
76.00
36.28%
30.09%
16.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seq primer: TF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSS.
                                                                                              11
                                                                                                                                                                                                                                                                                                                                                                                                                                                               533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CG270748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                                                                                               g
                                                                                                                                                                            ò
                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OGWIB33TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0583E18, genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Spermatophyta; Bubryophyta; Embryophyta; Tracheophyta; Byermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 934)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraecr, C.M., Budinan, M.A., Bedell, J.A., Robling, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
                                                                                                                                                                                                207 ACGAGGAAGTICGICAACTICCCGGAGATCGICGGAGGGGGTCCAGAAGGCTGGGTTCGAG 266
                                                                                                                                                                                                                                                                            267 GTCATCCCGATCGAGCCGCGTAGAGACCTCAGCGTGGAGGACTTCGCCCCGGACCGTCGAC 326
                                                                                                                                                                                                                                                                                                                                                            357
                                    87 récecadecidicaccareccaracaadeageceaacadeaacerecegeagerececeic 146
                                                                                                                  147 GGGGCCCCAGCTGAGCAGCTGCAGCGGAAGCCACGACTGATGCTTATCAACCGTGGCCGC 206
                                                                             53
                                                                                                                                                           36
                                                                                                                                                                                                                                       53
                                                                                                                                                                                                                                                                                                                   73
-----AlaPro 16
                                                                                                                                                                                                                                   53 rThrLeuValTyrAsnCygAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh
                                                                           17 Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref="taxon:4577"
/clone="ZMMBMa058BE18"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBGSK-; Site_1: HincII; 0.7-1.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93
44
93
44
93
                                                                                                                                                                                                                                                                                                                                                          327 rcgrece----Acerecrearesecececeses
                                                                                                                                                                                                                                                                                                                                                                                                                      SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                   eGlySerValAsnProAlaThrProThrAlaAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /mol_type="genomic DNA"
/strain="B73"
                                                                                                                                                         30 Thr-GlySerSerSerValAsp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .934
/organism="Zea mays"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x CG270741 (1-934)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CG270741.1 GI:34182882
GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unpublished (2002)
Other_GSSs: OGWIB33TV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seq primer: TR
Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Fax: 301-838-0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                        358
```

RESULT 14 CG270741/c LOCUS DEFINITION

ò 요 ò

VERSION KEYWORDS SOURCE ORGANISM

ACCESSION

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

FEATURES

Query Match: DB:

Pred. No.:

ORIGIN

US-09-917-376-4 (1-89) x CG270748 (1-1001)

```
598 GTCATCCCGATCGAGCCGCGTAGAGCCTCAGCGTGGAGGACTTCGCCCGGACCGTCGAC 657
                889 -----
                                                                                                                                                                    37 ------LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSe 53
                                                                                                                                                                                                                             53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh 73
2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer------AlaPro 16
                                                     17 Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn------ 29
                                                                                                             30 Thr-GlySerSerSerValAsp------36
                                                                                                                                                                                                                                                         658 TCGTGCG-----ACGTGCTCATGGGCGCGCACGGGG----
                                                                                                                                                                                                                                                                                                     73 eGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                         Search completed: May 12, 2004, 11:39:00 Job time: 840.94 secs
                      С
                                                     ઠે
                                                                                 g
                                                                                                             δ
                                                                                                                                       D
                                                                                                                                                                ò
                                                                                                                                                                                               g
                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                  ò
```

Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli Appli

Perfect score: Sequence:

Searched:

OM protein

Run on:

```
Fabrics Using Truncated
                                                                                                                                                                                                                                        Sequence
Sequence
Sequence
Sequence
                                                                                                                                        Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                             Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                       Sequence
Sequence
Sequence
Sequence
                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                    Sequence
                                                                    Sequence
                                                                                                Sequence
                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
CORRESPONDENCE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
OF INVENTION: Compositions and Methods for
Treating Cellulose Containing
Cellulase Enzyme Compositions
                                     US-08-460-462-3
US-08-460-457-1
US-08-460-458-3
US-08-460-458-3
US-08-460-458-1
US-08-460-458-1
US-08-300-394A-1
US-08-330-394A-3
US-09-306-632-5
US-09-006-632-5
US-09-006-632-5
US-09-277-716-21
US-09-277-716-21
US-09-277-716-21
US-09-278-21
US-09-526-193A-21
US-09-526-193A-21
US-09-526-193A-20
US-09-335-586-3
US-09-526-193A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-722A-20
US-09-233-338-39
US-09-233-338-39
US-09-233-338-338-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: «UNKnown»
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: SEPtember 19, 1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                US-09-233-752A-19
                                                                                                                                                                                                                                                                                                                                                                                                                -09-402-036-19
                                                                                                                                                                                                                                                                                                                                                                                                                             -09-904-226-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Farrington, Graham K. Anderson, Paige Gibbs, Moreland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 46, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Bak, Mary E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES:
                                                                                                                                                                                              499
1146
11146
8107
20284
25165
2241
2241
2241
2241
2241
2370
2370
2370
2370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 19477
US-09-136-574A-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                       75.5
72.5
70
                                                                                                                                                                                                                                                                                                              67.5
67.5
67.5
67.5
67.5
67.5
67.5
67.5
                                                                                                                                                                                                                                                                                                67.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1
 Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh.
-Q=/cgn2 1/0875PTO spool/1809917376/runat 11052004 114536 28265/app_guery.fasta_1.1429
-Q=/cgn2 1/0875PTO spool/1809917376/runat 11052004 114536 28265/app_guery.fasta_1.1429
-DB=188ued_Patents_NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL. -OUTFWT=pro-NOFM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15
-USFR=US9917376 @GGN 1 1.116 @runat 11052004 114536 28265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
Appli
Appli
Appli
, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Appli
Appli
Appli
                                                                                                May 12, 2004, 04:43:36 ; Search time 22.2257 Seconds (without alignments) 2222.227 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11, 8
Sequence 9, Ag
Sequence 11, 7
Sequence 6, Ag
Sequence 5, Ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 46, 1
Sequence 2, Ag
Sequence 1, Ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1,
Sequence 3,
Sequence 9,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                       470
1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptGdata/2/ina/5A_COMB.seq:*
/cgn2_6/ptGdata/2/ina/5B_COMB.seq:*
/cgn2_6/ptGdata/2/ina/6A_COMB.seq:*
/cgn2_6/ptGdata/2/ina/6B_COMB.seq:*
/cgn2_6/ptGdata/2/ina/PCTUS_COMB.seq:*
/cgn2_6/ptGdata/2/ina/PCTUS_COMB.seq:*
                5.1.6
Compugen Ltd.
                                                                       nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-136-574A-46
US-09-136-574A-2
US-09-136-574A-1
US-07-862-588B-1
US-09-339-159B-3
US-09-198-956-9
US-09-198-956-9
US-09-674-531-11
US-09-670-141-9
US-07-662-588B-6
US-07-862-588B-6
US-07-862-588B-6
                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                   682709 segs, 277475446 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                GenCore version
Copyright (c) 1993 - 2004
                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                           BLOSUM62
Xgapop 10.0, Xgapext
Ygapop 10.0, Ygapext
Fgapop 6.0, Fgapext
Delop 6.0, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Issued_Patents_NA:*
                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             B
                                                                                                                                                        US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2029
6416
11707
2977
1438
1482
1482
1482
1482
1482
1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41.4
41.0
335.1
335.1
335.1
335.1
255.6
25.6
                                                                                                                                                                                                              Scoring table:
```

Score

Result

Database :

```
APPLICATION NUMBER: US/09/136,574A
                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                               JS-09-136-574A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||||||| ::::|||
178 TGGGACAGATAGGGGAAGCAATGTGACATTCAATTTTGTGAAGCTTAGCAGCGGAGTG 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                            ValSerGlyGlyValLysValGlnTyrLysAsnAspAspAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                               41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House STATE: PA
                                                                                                                                                                                                                                                              2029
38
115
1
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Anderson, Graham K. Anderson, Paige Gibbs, Moreland Bergquist, Peter Daniels, Roy Morgan, Hugh W. Williams, Diane P. TITLE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
               REFERENCE/DOCKET NUMBER: 1997US001/CIP
                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x US-09-136-574A-46 (1-2029)
                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                             SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         238 AGTGGAGCGGATTATTACCTGGAG 261
                           TELECOMMUNICATION INFORMATION:
TELEPRAK: 215-540-9200
TELER: «Unknown-
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 2029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81 ProThrAlaAspThrTyrLeuGln 88
REGISTRATION NUMBER: 31,215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 19477
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                            195.50
60.23%
43.18%
41.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                             US-09-136-574A-46
                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-136-574A-2
                                                                                                                                                                                                                                                              No.:
                                                                                                                                                                                                                                                                Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         品
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
```

```
Containing Fabrics Using Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |||||||::||| ||||||||
2785 TACTGGTACACAGTGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2665 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAGGCCG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6416
38
14
32
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Treating Cellulose Containing
Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TILE OF INVENTION: Compositions and Methods for
                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x US-09-136-574A-2 (1-6416)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Farrington, Graham K.
Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
               CLASSIFICATION: <Unknown>
DATE: 19-Aug-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 1, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                LENGTH: 6416 base pairs
                                                                                                                                                                                                                            TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                         TYPE: nucleic acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2902 GATTATTACCTGGAG 2916
                                                                                                                                                                                                                                                  TELEX: «Unknown»
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.26e-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            194.50
61.18%
44.71%
41.38%
                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQUENCES:
```

```
144 CGACTGCAGTACAGAGCGGCCGATACAAATGCAGCCGACAACCAGATCAAGCCGTCCTTC 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26 GlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrp 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 LysvalGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46 PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2977
36
19
25
3
                                                                                                                                   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
                                                                                                                                                                                                                                            FILING DATE: 19920727
CLASSIFICATION: 435
PRIOR APPLICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: 221500, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFAX: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4 (1-89) x US-07-862-588B-1 (1-2977)
  405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INFORMATION FOR SEQ ID NO: 1: SEQUENCE CHARACTERISTICS: LENGTH: 2977 base pairs TYPE: NUCLEIC ACID STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: DNA (genomic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.05e-14
182.50
66.27%
43.37%
38.83%
                                                                                ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGINAL SOURCE:
ORGANISM: Bacillus
STRAIN: NCIMB 40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 375 TACGTGGAG 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86 TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION: 677..2
OTHER INFORMATION
                                      STATE: New York
COUNTRY: U.S.A.
ZIP: 10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                       New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
TOPOLOGY: lir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-07-862-588B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ଚ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4038 GGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCGAGCACACAAGTTCTATAAAGGCCG 4097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4158 TACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 4214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/07862588B
Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
                                                                                                                                        CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: «Unknown»
PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: BAK, MAY B.
REGISTRATION NUMBER: 1997US001/CIP
TELECHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-4 (1-89) x US-09-136-574A-1 (1-11707)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                TELEX: <UNKNOWN>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11707 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
                                                                              MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4275 GÁTTATTÁCTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.15e-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          84 AspThrTyrLeuGln 88
STATE: PA
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     192.50
62.35%
43.53%
40.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-07-862-588B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
```

à

ò

ò 셤 65

85

```
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-198-955A-11
; SEQ ID NO 9
; LENGTH: 1482
                                                                                                                                     Alignment Scores
                                                                                          US-09-198-956-9
                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1000 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAACTC 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1117 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 9, Application US/09198956
| Sequence 9, Application US/09198956
| Patent No. 6165769
| GENERAL INFORMATION:
| APPLICANT: Andersen, Lene N. APPLICANT: Schulein, Martin
| APPLICANT: Bjornvad, Mads E. APPLICANT: Bjornvad, Mads E. APPLICANT: Bjornvad, Mads E. APPLICANT: Schnorr, Kirk
| APPLICANT: Schnorr, Kirk
| APPLICANT: Schnorr, Kirk
| TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus FILE REFERENCE: 5377.200-US
| CURRENT APPLICATION NUMBER: US/09/198,956 |
| CURRENT APPLICATION NUMBER: US/09/198,956 |
| EARLIER FILING DATE: 1997-11-24 |
| EARLIER FILING DATE: 1997-11-24 |
| EARLIER FILING DATE: 1997-11-24 |
| EARLIER FILING DATE: 1997-11-25 |
| NUMBER OF SEQ ID NOS: 26 |
| NUMBER OF SEQ ID NOS: 26 |
| NUMBER OF SEXENCE FASTERO FOR Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1438
23
23
10
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-4 (1-89) x US-09-339-159B-3 (1-1438)
                                                                                                                                                                                                                      TITLE OF INVENTION: No. 6566114el Mannanases FILE REFERENCE: 5440_204-US CURRENT APPLICATION UNDBER: US/09/339,159B NUMBER PILING DATE: 1999-06-24 NUMBER OF SEQ ID NOS: 55 SOFTWARE: Patentin version 3.1
                                       Sequence 3, Application US/09339159B Patent No. 6566114 GENERAL INFORMATION APPLICANT: Kauppinen, Markus APPLICANT: Schulein, Martin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8.98e-13
                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: DNA ORGANISM: Bacillus sp. 1633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       165.00
59.79%
36.08%
35.11%
                                                                                                                                                      APPLICANT: Schnorr, Kirk
APPLICANT: Andersen, Lene
APPLICANT: Bjornvad, Mads
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 TrpAlaAlaMet~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 3
LENGTH: 1438
                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-339-159B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 6
US-09-198-956-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                              40
                                                                                                                                                                                                                                                                                                        1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --- GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                              21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                         1482
35
23
29
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Sorne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kongblak, Lars
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILE REFERENCE: 53798.200-02
CURRENT FILING DATE: 1998-11-24
                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                             US-09-917-376-4 (1-89) x US-09-198-956-9 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION UNMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR PELING DATE: 1997-11-24
PRIOR PELING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR PILING DATE: 1997-12-02
PRIOR PILING DATE: 1997-12-02
PRIOR PILING DATE: 1997-12-02
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-11-02
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASTSEQ for Windows Version 4
SEQ ID NO 11
LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 11, Application US/09198955A Patent No. 6187580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TYPE: DNA
; ORGANISM: Clostridium thermocellum
Ug-09-198-955A-11
TYPE: DNA ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 TrpAlaAlaMet------
                                                                                                       9.37e-13
165.00
59.79%
36.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Andersen, Lene N.
Schulein, Martin
```

8

ò

ò

ð 셤 ð

```
1168 CATGCTGCAATAATCGGCAGTAACGGCCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         991 GTATCAGGCAATTTGAAGGTTGAATTCTACAACAGCAATCCTTCAGATACTACTAACTCA 1050
                                                                                                                                                                                                                                                                                                                                             21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                                     --- GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                       41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Accounts in the No.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels Erik K.
APPLICANT: Bernad, Mads E.
APPLICANT: Bernad, Mads E.
APPLICANT: Schnort. Kirk
TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-08
CURRENT APPLICATION NUMBER: US/09/670,141
FILE REPERENCE: 5377.200-09
FRIOR APPLICATION NUMBER: 1991-124
PRIOR FILING DATE: 1999-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1482
33
23
29
10
  233
23
23
20
20
20
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                     US-09-917-376-4 (1-89) x US-09-694-531-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-4 (1-89) x US-09-670-141-9 (1-1482)
                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                       61 TrpAlaAlaMet------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 9, Application US/09670141
Patent No. 6429000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: DNA ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9.37e-13
165.00
59.79%
36.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Andersen, Lene N.
165.00
59.79%
36.08%
35.11%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-670-141-9
                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                           1111 ACATTGAGATATTATTATACAGTAGACGGACAGAAGATCAGACCTTCTGG----TGTGAC 1167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1168 CATGCTGCAATAATCGGCAGTAACGGCAACTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                      1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                               --- GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                  Thr ValArg Tyr TrpPheThr Arg Asp Gly Gly Ser Ser Thr Leu Val Tyr Asn CysAsp
                                                                                                                                                                                                                                                                                                    21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                     1482
35
23
29
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1482
                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6368843el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/694,531
CURRENT FILING DATE: 2000-10-23
                                                                                                                                                                               US-09-917-376-4 (1-89) x US-09-198-955A-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CURKENT FILLING DATE: 2000-120-62
PRIOR FILING DATE: 1998-11-24
PRIOR PLING DATE: 1998-11-24
PRIOR PLING DATE: 1998-11-24
PRIOR PLING DATE: 1997-11-24
PRIOR PLING DATE: 1997-11-24
PRIOR PRILING DATE: 1997-11-24
PRIOR PELING DATE: 1997-11-24
PRIOR PELING DATE: 1997-11-20
PRIOR PELING DATE: 1997-12-02
PRIOR PELING DATE: 1997-12-02
PRIOR PELING DATE: 1997-12-02
PRIOR PELING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-05-06
PRIOR PILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASELSEQ FOR WINDOWS Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 TrpAlaAlaMet------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: DNA ORGANISM: Clostridium thermocellum US-09-694-531-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 11, Application US/09694531
                                     9.37e-13
165.00
59.79%
36.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Schulein, Martin
Lange, Niels E.
Bjornvad, Mads E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Andersen, Lene N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Moller, Soren
                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                    41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               72
                                         .. oN ..
                                                               Score:
```

```
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                       No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
                                                           -----GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                            1228 ACATTIGIAAAATGAGIICCICAACAAATAACGCAGACACCIACCIIGAA 1278
                                                                                                                                           72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                    APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1624
28
18
36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
PPLICATION UMBER: US/07/862,588B
FILING DATE: 19920727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              62nd floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 19920727
FILING DATE: 19920727
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, El
REGISTRATION NUMBER: 30,335 / 33,728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
Compurer: IBM PC Compartible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
                                                                                                                                                                                                                                                          Sequence 6, Application US/07862588B Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                                                                                                              405 Lexington Avenue,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELECOMMUNICATION INPORMATION
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ 1D NO: 6: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacillus lautus
                                                             61 TrpAlaAlaMet-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 1624 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120.50
54.128
32.948
25.648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN: NCIMB 40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; OTHER INFORMATION US-07-862-588B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: CDS
LOCATION: 30...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE:
ORGANISM: Bac:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                       RESULT 11
US-07-862-588B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                  요
                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1051 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAACTC 1110
                                                               1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                           ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                     ---GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1482
35
23
10
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Markus S.
TITLE OF INVENTION: No. 6677147e] Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION WUMBER: US/09/198,955
PRIOR APPLICATION NUMBER: US/09/198,955
PRIOR PILING DATE: 1998-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR PILING DATE: 1997-12-02
PRIOR PLING DATE: 1997-12-02
PRIOR PLING DATE: 1997-12-02
PRIOR PLING DATE: 1999-12-02
PRIOR PLING DATE: 1998-05-06
PRIOR PLING DATE: 1998-05-06
PRIOR PLING DATE: 1998-05-06
PRIOR PLING DATE: 1998-05-06
PRIOR PLING DATE: 1998-01-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x US-10-072-152-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: DNA ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                    Sequence 11, Application US/10072152
Patent No. 6677147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.37e-13
                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Andersen, Lene N. APPLICANT: Schulein, Martin APPLICANT: Lange, Niels E.
                                                                                                                                                                                                                                                                                                                                                                                                                                             ornvad, Mads E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glad, Sanne O. S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         165.00
59.79%
36.08%
35.11%
                                                                                                                     61 TrpAlaAlaMet---
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Moller, Soren
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1482
                                                                                                                                                                                                                                                                               RESULT 10
US-10-072-152-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-10-072-152-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO 11
                                             41
                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                       ò
                                                                            g
                                                                                                                     ò
                                                                                                                                                          셤
                                                                                                                                                                                                   à
                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
NCIMB 40250
                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10036
                                                                                                                                                                                                                Alignment Scores:
                                        NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-048-164A-1
                                                                                                                                                                                                                                      ۰:
و
                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                  ||||||| :::|||||| :::
1383 AAGCCTTCCTTTAACATCAAAGGTACAACGGTACAACCCTGTTAACCTGAGCGGCCTCAAG 1442
                                                                                                                                                                                                                                                            1500 GCGCAAATCGGCCGAACGAATGTTCTGCTGGCATTC----GCTAACTTTACCGGGAGT 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor CITY: New York COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                              62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                        SerglyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                      LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
                                      US-09-917-376-4 (1-89) x US-07-862-588B-6 (1-1624)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION NUMBER: PCT/DK91/00013
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, El
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     335 / 33,728
3425.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-07-862-588B-5
; Sequence 5, Application US/07862588B
; Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 34
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                              1554 AATACGGATACTTAC 1568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: DNA (genomic)
FRACMENT TYPE: N-terminal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                        82 ThrAlaAspThrTyr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1775 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUCLEIC ACID
EDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10017
                                                                                                            임
                                                                                                                                                      ò
                                                                                                                                                                                          엄
                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                  δ
```

```
1443 CTTNNNNNNNNNNNNAAAAAGAC---GGACCTGCGGATATGAGCTGCTCGATCGACTGG 1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1500 GCGCAAATCGGCCGAACGAATGTTCTGCTGGCATTC-----GCTAACTTTACCGGGAGT 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                /evidence= EXPERIMENTAL
/transl_except= (pos: 1446 .. 1458, aa:
OTR)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,164A
FILING DATE: 14-APR-1993
CLASSIFTAMENTAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shoiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x US-07-862-588B-5 (1-1775)
                                           DD: experimental
/partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1, Application US/08048164A Patent No. 5496934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 780
FELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1554 AATACGGATÁCTTÁC 1568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AFELLING DATE: 14-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leglie
NAME: Misrock, S. Leglie
LOCATION: 30..(1625.1775)
IDENTIFICATION METHOD: ex
                                                                                                                                                                                                                                                                                                                                                                                                                                1.32e-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 ThrAlaAspThrTyr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120.50
54.12%
32.94%
25.64%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                       CTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
```

```
JS-08-048-164A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-460-462-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY:
                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 15
                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||||
184 GCTGGTGCATTATTAGGAAATAGCTATGTTGATAACACTAGCAAAGTGACAGCAAACTTC 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                              SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 LysproGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 AlaAlaMetGlyCysGlyAsn-----11eArgAlaSerPhe 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        244 GTTAAAGAAACAAGCCCA---ACATCAACCTATGATACATAGTTGAA 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                          4422214
8822414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x US-08-048-164A-1 (1-486)
                                                                                                                                                                                                                                                                                            Matches:
                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US/08/048,164A FILING DATE: 14-APR-1993 CLASSIFICATION: 435 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/WS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
US-08-048-164A-3/c
; Sequence 3, Application US/08048164A
; Patent No. 5496934
               TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
(212) 790-9090
                                                                                                                                                                                                                                                                            .64e-06
                                                                                                                                                                                                                                                                                          1114.00
52.58%
29.90%
24.26%
                                                                                  LENGTH: 486 base pairs
TYPE: nucleic acid
                                                                                                                       double
                                                                                                                                     unknown
                                                                                                                                     TOPOLOGY: unknow MOLECULE TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                      NAME/KEY: CDS
                                                                                                     TYPE: nucleic
STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 10036
 TELEPHONE:
                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                        LOCATION:
                                                                                                                                                                                                                       US-08-048-164A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42
                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
303 GCTGGTGCATTATTAGGAAATAGCTATGTTGATAACACTAGCAAAGTGACAGCAAACTTC 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Goldstein, Marc A.
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----IleArgAlaSerPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       243 GTTAAAGAAACAGCCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                  486
222
481
124
124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-4 (1-89) x US-08-048-164A-3 (1-486)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                In 1155 Avenue of the Americas
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,462
FILING DATE: CONCURRENTLY herewith
                                            7809-003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62 AlaAlaMetGlyCysGlyAsn----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-
TELECOMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEPHONE: (212) 869-8864/9741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1, Application US/08460462
Patent No. 5670623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PENNIE & EDMONDS
                                                                                                                     TELEX: 66141 PENNIE
INPORMATION POR SEQ ID NO: 3
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
                                                                                                                                                                                                                                                                                                                                                       .64e-06
                                                                                                                                                                                                                                                                                                                                                                   114.00
52.58%
29.90%
24.26%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
                                                                                                                                                                                                                                                                  MOLECULE TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE:
                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE: Ne
COUNTRY:
ZIP: 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE:
```

```
|||||||||||::::::|||
127 GTTAGATATTATTACACAAGTGATGGTACACAAGGACAAACTTTCTGG---TGTGACCAT 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----IleArgAlaSerPhe 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  244 GTTAAAGAAACAGCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-4 (1-89) x US-08-460-462-1 (1-486)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14.APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leelie
REGISTRATION NUMBER: 18,872
REPERENCE/DOCKET NUMBER: 7809-006
TELEPHONE: (212) 790-9090
TELEPAX: (212) 869-8864/9741
TELEFAX: (512) 869-8864/9741
TELERAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPPLOGGY: unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: May 12, 2004, 15:15:25 Job time: 27.2257 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AlaAlaMetGlyCysGlyAsn-
                                                                                                                                                                                                                                                                                                                                                                                                                           1.64e-06
114.00
52.58%
29.90%
                                                                                                                                                                                                                                                                                         TOPOLOGY: unknown MOLECULE TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                           NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                              US-08-460-462-1
                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
```

Appli Appli Appli Appli Appli

Perfect score: Sequence:

Run on:

Searched:

```
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 21, Appl
Sequence 22, App
Sequence 22, App
Sequence 22, App
Sequence 21, Appl
Sequence 21, Appl
Sequence 1, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 18, Appl
                                                                                                                                                                               Sequence 1, Appli
Sequence 46838, A
Sequence 3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 9462, Āp
Sequence 3, Appli
                                                                                                                                                                                                                                                                         Sequence 11, App.
Sequence 11, App.
Sequence 684, App
                                                                 Sequence 2, App
Sequence 2, App
Sequence 2, App
Sequence 550, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 85, 7
Sequence 10, 7
Sequence 21, 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 21
Sequence 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 2, Application US/09917378

Publication No. US20030119093A1

GENERAL INFORMATION:

APPLICANT: DING, SHIL-YOU

APPLICANT: DING, SHIL-YOU

APPLICANT: DING, SHIL-YOU

APPLICANT: VINZANT: TODD B.

APPLICANT: VINZANT: TODD B.

APPLICANT: WINCHER, STEPHEN R.

APPLICANT: HIMMEL, MICHAEL E.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION UNDERS: US/09/917,378

CURRENT FILING DATE: 2001-07-28

SOFTWARE: PATENTIN Ver. 2.1

SEQ ID NO 2

LENGTH: 2289
15 US-10-156-761-1

9 US-09-764-860-85

15 US-10-074-095-85

16 US-10-212-872-85

10 US-09-984-827-10

13 US-10-617-334-21

16 US-10-452-510-21

13 US-10-425-114-9462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-822-268A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-2
           2289
3687
3687
2869
2869
2223
9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1146
3147
96597
2091
2091
3396
4041
4353
4914
4914
4914
5714
5714
5714
5773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2402
48667
                                                                                                                                                                                                                                                                                                       1482
1314
4059
4056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20284
20284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                            US-09-917-378-2
              466
466
466
466
466
466
1190
1190
1190
1119
1119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Command line parameters:
-MODEL=frame+ p2n.model_DEV=xlh
-MODEL=frame+ p2n.model_DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09917376/runat_11052004_114537_28336/app_query.fasta_1.1429
-Q=/cgn2_1/USPTO_spool/US09917376/runat_1052004_114537_28336/app_query.fasta_1.1429
-DB=Published Applications NA -QFWT=fastap -SUFFIx=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -RND=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LCCAL -OUTFWT=pct -NORM=ext -HBAPGTZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09917376 @CCN 1 1 809 @runat 11052004 114537_28336
-LOOFLCG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10-YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                                                  May 12, 2004, 09:26:52 ; Search time 136.654 Seconds (without alignments) 2950.286 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                    470
1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_BEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
              GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2941586 segs, 2264995651 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                         0.5
7.0
7.0
                                                                                                                                                                                                                                                                                                                                                                                                                     BLOSUM62
XGapop 10.0, Xgapext
Ygapop 10.0, Ygapext
Fgapop 6.0, Fgapext
Delop 6.0, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Score Match Length DB
                                                                                                                                                                                                                                                                                                          US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                         Scoring table:
```

Jatabase

Result

```
RESULT 4
US-09-911-376-2
Sequence 2, Application US/09917376
; Publication No. US20040038334A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1.13e-55
466.00
100.00%
98.86%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-383-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1990
                             41
                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                           원
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                      1810 ATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGTCGGTGGATTTGTCGACGGTG 1869
                                                                                                                                                                                                                                                                                                                                                               1423 ATCAAACCGGGTCTCCAGTTGGTGAATACGGGGTCGTCGTCGTCGATGGATTTGTCGACGGTG 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1483 ACGGGGGACTGCTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1543 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                                                                                                                                                                                 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                               41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TrpAlaAlaMetGlyCysGlyAsnileArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2, Application US/09917384
; Bequence 2, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
    APPLICANT: DING, SHI-YOU
    APPLICANT: DECKER, WILLIAM S.
    APPLICANT: DECKER, STEPHEN R.
    APPLICANT: HIMMEL, MICHAEL E.
    TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
    TITLE OF INVENTION: CELLULOLYTICUS
    FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 3687
; TYDE: DATE: DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM DATE: DECKER, WILLIAM 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER INFORMATION: Description of Artificial Sequence: Segment OTHER INFORMATION: GuxA
                      2289
888
0
0
0
                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-4 (1-89) x US-09-917-384-2 (1-3687)
                                                                                                                                                                                         US-09-917-376-4 (1-89) x US-09-917-378-2 (1-2289)
                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1603 CCGACGGCGACACCTACCTGCAG 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: DNA
ORGANISM: Artificial Sequence
                 2.21e-56
469.00
100.00%
100.00%
99.79%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.13e-55
466.00
100.00%
98.86%
99.15%
                                                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-384-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; OTHER INFORT
US-09-917-384-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
                      . No. :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 8 ×
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGACGTCGACACTGGTGACACTGTGAC. 1929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGACACTGTGAC 1929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1930 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCTCGTTCGGCTCGGTGAACCCGGCGACG 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                  TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-383-2

| Sequence 2, Application US/09917383
| Sequence 2, Application US/09917383 |
| Publication No. US20030104522A1 |
| GENERAL INFORMATION: |
| APPLICANT: DING, SHI-YOU |
| APPLICANT: DECKER, STEPHEN R. |
| APPLICANT: UNDARY, TOOD B. |
| APPLICANT: HEMMEL, MICHAEL E. |
| TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVENTION: CELLULOLYTICUS |
| TITLE OF INVEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-4 (1-89) x US-09-917-383-2 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCGACGGCGGACACCTACCTGCAG 2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                 81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81 ProThrAlaAspThrTyrLeuGln 88
```

```
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c,
US-10-155-400-2
                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-156-761-550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , NAME/KEY: CDS
, LOCATION: (1)
US-10-156-761-550
                                                                                                   Alignment Scores:
                                                                                                                                                                                                Query Match:
                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTGGATTTGTCGACGGTG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2785 TGGGCGGCGATCGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGTGAACCCGGCGACG 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TrpAlaAlaMetClyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
          APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197-440501
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT PILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/10155400

Publication No. US2003010898A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: DING, WILLIAM S.

APPLICANT: TINAMY, TODD B.

APPLICANT: HIMMEL, MICHAELE S.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REPRESENCE: NRED. 01-36A

CURRENT APPLICATION NUMBER: US/10/155,400

CURRENT PILING DATE: 2002-10-22

NUMBER: OF SEQ ID NOS: 7

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                               other or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x US-09-917-376-2 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2845 CCGACGCGGACACCTACCTGCAG 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: DNA
ORGANISM: Acidothermus cellulolyticus
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                             TYPE: DNA ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                        ; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g,
US-09-917-376-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.98e-55
462.00
100.00$
97.73$
98.30$
                                                                                                                                                                                                                                                                                                                                    NAME/KEY: modified_base
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
2785 TGGGCGGCGATCGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTCGATTTGTCGACGGTG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2605 Grerceseresesreaaserecasraraasaaraaraaresereseeseserearaareas 2664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2223
49
13
25
0
                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
other or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-4 (1-89) x US-10-155-400-2 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECOND TO SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND SECOND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: DNA ORGANISM: Streptomyces avermitilis
       ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.69e-26
257.00
71.26%
56.32%
54.68%
                                                                                                                                          2.98e-55
462.00
100.00$
97.73$
   'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .. (2223)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
DB:
```

```
APPLICANT: Cao, Yongwei
APPLICANT: Cao, Yongwei
APPLICANT: Cao, Yongwei
APPLICANT: Slater, Seeven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Goldman, Barry S.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROPERTIES
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR PILING DATE: 2003-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 46838
                                                               706396 GCGGCCGTCGGCTGTTCCAACGTCAGCCTGAAGGTCGTACCCCTGACCACGCCCGTGCCG 706455
706336 GCCGGCTACTACTTCACCCGGGACAGCGGCTCGCCCACCGTGAACGCCTGGTGCGACTAC 706395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||||||||::
|204 TACTGGTATAAAGGGAAAAAAAAGGCCAAAAC-----TTTGACTGTGACTACGCG 1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1255 CAGATTGGATGCGGCAATGTGACACACAGTTTGTGACGTTGCATAAACCAAAGCAAGGT 1314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 GlyValLysValGlnTyrLysAsnAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TyrTrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAla
                                   62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1527
35
21
26
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-4 (1-89) x US-10-369-493-46838 (1-1527)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                      706456 GGAGCCGACCTACCTCGAA 706476
                                                                                                                                                                                                         S-10-369-493-46838
Sequence 46838, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
                                                                                                                82 ThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 3, Application US/10172054
Publication No. US20030203466A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.02e-17
190.00
65.12%
40.70%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Kauppinen, Markus
APPLICANT: Schulein, Martin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: DNA ORGANISM: Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Schnorr, Kirk
Andersen, Lene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JS-10-369-493-46838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: APPLICANT: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1951 GCGGCCGTCGGCTGTTCCAACGTCAGCCTGAAGGTCGTACCCCTGACCACGCCCGTGCCG 2010
                                                                                                                41
                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                       21
                                                                                                                                                                                                                                                                   62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                       2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                          42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                              22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-4 (1-89) x US-10-156-761-1 (1-9025608)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    other or unknown
US-09-917-376-4 (1-89) x US-10-156-761-550 (1-2223)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT APPLICATION NUMBER: US/29
PRIOR APPLICATION NUMBER: US/201-204089
PRIOR FILING DATE: 2001-05-30
PRIOR PRILING DATE: 2001-05-30
PRIOR FILING DATE: 2001-06-02
NUMBER OF SEQ ID NOS: 15109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                  2011 GGAGCCGACGCCTACCTCGAA 2031
                                                                                                                                                                                                                                                                                                                                              82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Streptomyces avermitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1, Application US/10156761 Publication No. US20030119018A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION: a, t, c, g,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.72e-21
257.00
71.26%
56.32%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: OMURA, SATOSHI
APPLICANT: IREBA, HARUO
APPLICANT: ISHIKAWA, UUN
APPLICANT: HORIKAWA, HIROSHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY: misc_feature LOCATION: (4187715)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OTHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            유
                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                      셤
                                                                   g
                                                                                                              8
                                                                                                                                                                                          ò
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
```

```
Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-072-152-11
                                                                                                                                                                                                      US-10-655-433-11
                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1000 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAACTC 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        940 GTATCAGGCAATTTGAAGGTTGAATTCTACAACAGCAATCCTTCAGATACTACTAACTCA 999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Andersen, Lene N.
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Knogsbak, Lars
TITLE OF INVENTION: No. US20040067572Alel Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/10/655,433
CURRENT APPLICATION NUMBER: US/0.000-004
                                                                                                                                                                                                                                                                                                                1438
35
23
29
10
                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4 (1-89) x US-10-372-054-3 (1-1438)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CURKENT FILING DATE: 2003-09-04
PRIOR APPLICATION NUMBER: US/09/198,955A
PRIOR FILING DATE: 1998-11-24
PRIOR PILING DATE: 1998-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-05
PRIOR FILING DATE: 1997-11-05
PRIOR FILING DATE: 1997-11-05
APPLICANT: Bjornvad, Mads
TITLE OF INVENTION: Novel Mannanases
FILE REFERENCE: 5440.204-US
CURRENT APPLICATION NUMBER: US/10/372,054
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US/09/339,159B
PRIOR FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 55
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 1438
TYPE: DNA
ORGANISM: Bacillus sp. I633
US-10-372-054-3
                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 TrpAlaAlaMet------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 11, Application US/10655433
Publication No. US20040067572A1
                                                                                                                                                                                                                                                                                                     2.72e-13
165.00
59.79$
36.08$
35.11$
                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
US-10-655-433-11
                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                    ..
9
                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
|||||||:::
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         991 GTATCAGGCAATTTGAAGGTTGAATTCTACAACAGCAATCCTTCAGATACTACTAACTCA 1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ||||:::|||||||:::::|||| |||||||
1111 ACATTGAGATATTATTACAGTAGAGGACGACGAAAGATCAGACCTTCTGG---TGTGAC 1167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------GlyCysGlyAsnIleArgAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 ileLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Schnorr. Kirk
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr. Kirk
APPLICANT: Schnorr. Kirk
APPLICANT: Schnorr. Markus S.
APPLICANT: Markin Schnorr. Markus S.
CURRENT APPLICATION NUMBER: US/10/072,152
CURRENT FILING DATE: 2002-02-07
                                                                                                                                                                                                                                                                                                                                                         1482
335
23
10
10
                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89) x US-10-655-433-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CUKKENI FILING DATE: 2002-02-07

PRIOR APPLICATION NUMBER: US/09/198,955

PRIOR FILING DATE: 1998-11-24

PRIOR PELING DATE: 1998-11-24

PRIOR PELING DATE: 1997-11-24

PRIOR PELING DATE: 1997-11-24

PRIOR PILING DATE: 1997-11-24

PRIOR FILING DATE: 1997-11-24

PRIOR FILING DATE: 1997-12-02

PRIOR FILING DATE: 1997-12-02

PRIOR PILING DATE: 1997-12-02

PRIOR PILING DATE: 1997-12-02

PRIOR PILING DATE: 1997-12-02

PRIOR APPLICATION NUMBER: 09/057,684

PRIOR APPLICATION NUMBER: 09/073,684

PRIOR APPLICATION NUMBER: 09/073,684
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR PAPLICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: RateSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TrpAlaAlaMet------
                                                                                                                                                                                                                         TYPE: DNA ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 11, Application US/10072152
Publication No. US20020142438A1
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
                                                                                                                                                                                                                                                                                                                                                         2.84e-13
165.00
59.79%
36.08%
35.11%
                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
```

83

44

24

```
878 ATAGCAGTACAATACAGAGCGGGGACAACAATGTAAACGGCAACCAATCCGCCCTCAG 937
                                                                                                                                 45 TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
 5 ValLysValGinTyrLysAsnAsnAspSerAlaProGlyAspAsnGinIleLysProGly
                                                                                                                                                                                                      64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                  LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr
                                                                                                                                                                                                                                                                                                                                                                Sequence 1, Application US/09784554B
; Publication No. US20030032162A1
; GENERAL INFORMATION:
APPLICANT: Schnorr, Kirk
; APPLICANT: Schnorr, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-4 (1-89) x US-09-784-554B-1 (1-4059)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: DNA ORGANISM: Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3835 GATTATTATTIGGAA 3849
                                                                                                                                                                                                                                                                                                        1109 GACACGTATCTTGAA 1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                        84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2.5e-11
155.50
56.47%
31.76%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
US-09-784-554B-3
                                                                                                                                                                                                                                                                                                                                                          US-09-784-554B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-784-554B-1
                                                                  25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                            g
                                                                                                                                   ઠે
                                                                                                                                                                  요
                                                                                                                                                                                                      à
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                            |||||||||| :::|||||||:::::
991 GTATCAGGCAATTGAAGGTTGAATTCTACAACACAATCCTTCAGATACTACTAACTCA 1050
                                                                                                                                                                                                                                                                                                                                                                                                            |||:::|||||||::::::|||
| ACATTGAGATATTATTATACAGTAGACGGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                          IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Sequence 649, Application US/09974300
| Patent No. US20020146721A1
| GENERAL INNORMATION:
| APPLICANT: Berka, Randy M.
| APPLICANT: Clausen, Ib Groth
| TITLE OF INVENTION: Expression
| TITLE OF INVENTION: Expression
| FILE REFERENCE: 10085.500-US
| CURRENT APPLICATION NUMBER: US/09/974,300
| CURRENT APPLICATION NUMBER: 00/998
| PRIOR APPLICATION NUMBER: 00/680,598
| PRIOR PAPLICATION NUMBER: 60/279,526
| PRIOR PAPLICATION NUMBER: 60/279,526
| PRIOR PLICATION NUMBER: 60/279,526
| PRIOR PLICATION NUMBER: 00/279,526
| SOFTWARE: PastSEQ for Windows Version 4.0
| SOFTWARE: PastSEQ for Windows Version 4.0
                                                                                                                                                                    1482
35
23
29
10
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x US-09-974-300-684 (1-1314)
                                                                                                                                                                                                                                                                                        US-09-917-376-4 (1-89) x US-10-072-152-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TrpAlaAlaMet------
                                                                                TYPE: DNA ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3.34e-12
157.00
56.47%
37.65%
                                                                                                                                                                    2.84e~13
165.00
59.79%
36.08%
35.11%
                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
US-09-974-300-684
                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-974-300-684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                    LENGTH: 1482
                                                                                                                   US-10-072-152-11
                                                                                                                                                                                                                                                                                                                                                                                                                           1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1111
                                                                                                                                                                                                                                                                                                                                                                                        21
                                                                                                                                                                                                                                                                                                                                                                                                                                                           41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                   쉽
                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
3598 GGACTGCTGCTCCAATATCGCACAGCAGATAAGGTGAAACGATAATCACCTCAATCCG 3657
                                                                                                                                                                                                                                                                                                                                                                     ::: |||
3775 CTGAGCTGCTCAAAGCTGAAAGGTAAAATGGAGAAGGCTGCAACCGGTGCC 3834
                                                                                                                                                                                                                                                                                                                                                                                                                                       44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                   4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                     24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 3, Application US/09784554B
```

```
Search completed: May 12, 2004, 18:16:54
Job time : 817.654 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-282-122A-15299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              심
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3595 GGATTGGTGCTCCAGTATCGCACAGCGGATACAAATGTGAACGACAATCACTTGAACCCG 3654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3772 CTGAGCTGCTCGAAGCTGAATGGTAAGCTGGTTAAAATGGATAAAGCTGCAACCGGTGCT 3831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3715 TACTACTACACGATCGACGGTGACGTGAGCAG---ACATTCAACTGCGACTATGCGGTG 3771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms FILE REPERENCE: ELITRA.034A
CURRENT APPLICATION VIMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR PELICATION NUMBER: 60/191,078
PRIOR PLILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR PLILING DATE: 2000-05-23
PRIOR PLILING DATE: 2000-05-23
PRIOR PLILING DATE: 2000-05-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAshGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArg
                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-4 (1-89) x US-09-784-554B-3 (1-4056)
                                                       APPLICANT: Jorgensen, Per Lina
APPLICANT: Schulein, Martin
TILE OF INVENTION: PAMILY 44 XYLOGLUCANASES
FILE REFERENCE: 10017.200-US
CURRENT APPLICATION NUMBER: US/09/784,554B
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 15299, Application US/10282122A Publication No. US20040029129A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                    TYPE: DNA
ORGANISM: Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3832 GATTATTATTGGAA 3846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Malone, Cheryl
APPLICANT: Ohlsen, Kari
APPLICANT: Syskind, Judith
APPLICANT: Zyskind, Judith
APPLICANT: Tawick, John
APPLICANT: Trawick, John
APPLICANT: Trawick, John
APPLICANT: Yamamoto, Robert
APPLICANT: Yamamoto, Robert
APPLICANT: Yamamoto, Robert
                                                                                                                                                                                                                                                                                                                                                         1.29e-10
150.50
56.47%
31.76%
32.02%
Publication No. US20030032162A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84 AspThrTyrLeuGln 88
                                        APPLICANT: Schnorr, Kirk
                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-282-122A-15299
                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                          SEQ ID NO 3
LENGTH: 4056
                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                          US-09-784-554B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
```

```
41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AlaAlaMetGlyCysGly------AsnIleArgAlaSerPhe 73
                      PRIOR PILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR PILING DATE: 2000-11-27

PRIOR PILING DATE: 2000-11-27

PRIOR PILING DATE: 2000-12-22

PRIOR PILING DATE: 2000-12-22

PRIOR PILING DATE: 2000-12-22

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16

PRIOR PILING DATE: 2001-02-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4 (1-89) x US-10-282-122A-15299 (1-4452)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: DNA ORGANISM: Clostridium acetobutylicum
60/230,335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.04e-07
130.50
55.79%
31.58%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
```

```
GenCore version 5.1.6 (c) 1993 - 2004 Compugen Ltd.
                  Copyright
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 12:10:32 ; Search time 54 Seconds (without alignments) 465.680 Million cell updates/sec

US-09-917-376-4 89

ritle:

1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX Perfect score: Sequence:

OLIGO Gapop 60.0 , Gapext 60.0 Scoring table:

1586107 segs, 282547505 residues

Searched:

12 Total number of hits satisfying chosen parameters: 10 Word size

seq length: 0 seq length: 2000000000 Minimum DB Maximum DB Post-processing: Listing first 1000 summaries

Jatabase

A_Geneseq_29Jan04:* geneseqp1980s:* geneseqp1990s:*

geneseqp2003as:* geneseqp2003bs:* geneseqp2000s:* geneseqp2001s:* geneseqp2002s:*

geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	Add22925 Acidother	Add22924 Acidother	Abp73020 Amino aci	Abp73022 Amino aci	Abp73025 Amino aci	Abp73029 Amino aci	Abp71658 A. cellul	Abp71656 A. cellul	Abp73018 Amino aci	Abp73017 Amino aci	Abp73015 Amino aci	Add22921 Acidother
	ΠD	ADD22925	ADD22924	ABP73020	ABP73022	ABP73025	ABP73029	ABP71658	ABP71656	ABP73018	ABP73017	ABP73015	ADD22921
	Match Length DB	88 7	89 7	154 6	762 6	150 6	1228 6	153 6	1121 6	9 88	89 6	957 6	957 7
* Query	Match 1	98.9	98.9	98.9	98.9	93.3	93.3	79.8	79.8	40.4	40.4	40.4	40.4
	Score	88	88	88	88	83	83	71	71	36	36	36	36
Result	No.	н	7	٣	4	₂	•	7	æ	6	10	11	12

ALIGNMENTS

```
ADD22925 standard; protein; 88 AA.
                                                        (first entry)
                                                        15-JAN-2004
                                      ADD22925
RESULT 1
ADD22925
ID ADD2
XX
AC ADD2
XX
DT 15-J
```

Key Location/Qualifiers Misc-difference 89

Acidothermus cellulolyticus

```
New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to an isolated polynucleotide molecule encoding thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucleotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII CBD III #2.
                                   enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
Acidothermus cellulolyticus avicelase AviIII CBD III #2.
                                                                                                                                                                                                                                                                                                                                                              Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 16; SEQ ID NO 5; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                          Ding S, Adney WS, Vinzant TB,
                                                                                                                                                                                                 18-OCT-2002; 2002US-00155400.
                                                                                                                                                                                                                                    28-JUL-2001; 2001US-00917376
                                                                                      Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                   DING S.
ADNEY W S.
VINZANT T B.
                                                                                                                                                                                                                                                                   (DING/) DING S. (ADNE/) ADNEY W S. (VINZ/) VINZANT T B. (HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2003-810853/76.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 88 AA;
                                                                                                                           US2003108988-A1
                                                                                                                                                            12-JUN-2003.
```

៧

9 9 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD Gaps ; 0 enzyme, AviIII; cellulose reduction, agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase. 98.9%; Score 88; DB 7; Length 88; 100.0%; Pred. No. 1.9e-81; 0, Indels Acidothermus cellulolyticus avicelase AviIII CBD III #1. 100.0%; Pred. No. 100.0%; Pred. No. Mismatches 61 WAAMGCGNIRASFGSVNPATPTADTYLO 88 61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88 ADD22924 standard; protein; 89 AA. (first entry) 88; Conservative Local Similarity 15-JAN-2004 ADD22924; Query Match Best Loc Matches ADD22924 RESULT 硆 ð 셤 ò

ö

Ding

```
The present sequence represents a fragment of ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starking material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of
                                                                                                                               Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
feed; paper pulp; biofuel; mannase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 6; L
3.2e-81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino acid sequence of the ManA polypeptide.
                                                            Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98.9%; Score 88; DB 100.0%; Pred. No. 3.2 ive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant TB, Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABP73022 standard; protein; 762 AA
                                                            Adney WS, Vinzant TB,
                                                                                                                                                                                                                   Claim 8; Page 7; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88; Conservative
                      (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2003-248182/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding S, Adney WS,
                                                                                              WPI; 2003-248182/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ManA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; ABZ77633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 154 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO2003012110-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP73022;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                            Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ManA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                  New isolated thermal tolerant avicelase polymucleotide useful for detection of a polymucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to an isolated polynucleotide molecule encoding thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. To method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII CBD III #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Manh; mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar; food; feed; paper pulp; biofuel; mannase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino acid sequence of the ManA carbohydrate binding domain type III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.9e-81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 7;
                                                                                                                                                                                                                                                                                              Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Pred. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 88;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 16; SEQ ID NO 4; 29pp; English
                                                                                                                                                                                                                                                                                              Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABP73020 standard; peptide; 154
/label= Unknown
                                                                                                                  18-OCT-2002; 2002US-00155400.
                                                                                                                                                         28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                  WPI; 2003-810853/76
                                                                                                                                                                                                                                                     HIMMEL M E.
                                                                                                                                                                                                         ADNEY W S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                              Adney WS,
                                        US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO2003012110-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 89 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        03-JUN-2003
                                                                              12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP73020;
                                                                                                                                                                                                                                                        (HIMM/)
                                                                                                                                                                                                                                                                                            ຜ
                                                                                                                                                                                                               (ADNE/)
(VINZ/)
                                                                                                                                                                                               (DING/)
```

Matches

g

ઠે

g

ò

ö

Gaps . 0

Length 154; 0; Indels 9 9 for

reagent tools

or detecting GuxA polypeptide expression, as well as reag characterizing the molecular actions of GuxA polypeptides

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABP73023-28 represent peptide fragments of a GuxA polypeptide. GuxA is a thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, purp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                    polypoptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicallulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA.
                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GuxA; cellulase; glycoside hydrolase; enzyme; cellulose; sugar; biofuel;
detergent; pulp processing; paper processing; feed processing; textile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino acid sequence of the GuxA carbohydrate binding domain type III.
                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                      sequence represents ManA, a thermostable mannanase A
                                                                                                                                                                                                                                                                                                                                                              .;
0
                                                                                                                                                                                                                                                                                                                         Length 762;
                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Decker
                                                                                                                                                                                                                                                                                                                     98.9%; Score 88; DB 6; Le
100.0%; Pred. No. 1.4e-80;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 515 WAAMGCGNIRASFGSVNPATPTADTYLQ 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                 Claim 3; Page 18-19; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABP73025 standard; peptide; 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 6; Page 7; 47pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     03-JUN-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                            88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MIDE ) MIDWEST RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2003-239526/23.
                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Adney WS,
                                                                                                                                                                                                                                                                                    Sequence 762 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO2003012109-A1
                                                                                                                                                                                                                                                  polypeptides
                                                        The present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP73025;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ding S,
                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
```

```
ö
                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present sequence represents a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                              65
                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                               GuxA; cellulase; glycoside hydrolase; enzyme; cellulose; sugar; biofuel; detergent; pulp processing; paper processing; feed processing; textile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 KVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                              6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                              6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                              Gaps
                                                                                                              ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93.3%; Score 83; DB 6; Length 1228; 100.0%; Pred. No. 2.5e-75; ive 0; Mismatches 0; Indel8
                                                                              93.3%; Score 83; DB 6; Length 150; 100.0%; Pred. No. 3.7e-76;
                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                               Amino acid sequence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Himmel ME,
                                                                                                              0; Mismatches
                                                                                                                                                                                                                 88
                                                                                                                                                                                                                                              88
                                                                                                                                                                                                                                                                                                                               ABP73029 standard; protein; 1228 AA
                                                                                                                                                                                                               66 CGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                  66 CGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 3; Page 19; 47pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2003-239526/23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1228 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; ABZ77634.
                                                 Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO2003012109-A1.
                                                 Sequence 150
                                                                                                                                                                                                                                                                                                                                                                                              03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                ABP73029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                              Query Match
Best Local
                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                ន្តដូន
                                                                                                                                                ઠે
                                                                                                                                                                               셤
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                          셤
```

셤 ò 임

```
The invention relates to a thermal tolerant Guxl peptide from A. cellulolyticus. The Guxl exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type II. and a carbohydrate binding domain type II. and a carbohydrate binding domain type II. for section of cellulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial scale processing of cellulose to sugar for use in biofuel production. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                             Gux1; thermal tolerant; exoglucanase; glycoside hydrolase; cellulose; biofuel; detergent; pulp; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           scale processing of cellulose to sugar for use in biofuel prod
present sequence represents a A. cellulolyticus Guxl cellulase
                                                                                                                                                                                                                                                            1. .34
/note= "potential signal peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79.8%; Score 71; DB 100.0%; Pred. No. 3.6 ive 0; Mismatches
                                                                                                                                                                                                                                                                                                           'note= "mature protein"
                                                                                                                                                                                                                                                                                                                                           /note= "encoded by CG"
                  ABP71656 standard; protein; 1121 AA.
                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 3; Page 18-19; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Vinzant TB,
                                                                                                               A. cellulolyticus Gux1 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                              Acidothermus cellulolyticus.
                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           71; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2003-300494/29.
N-PSDB; ABZ76162.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                              cellulase; enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1121 AA;
                                                                                                                                                                                                                                                                                                                                                                       WO2003012095-A1.
                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                               29-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                        13-FEB-2003.
                                                 ABP71656;
                                                                                                                                                                                                                                                           Peptide
                                                                                                                                                                                                                                                                                            Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The invention relates to a thermal tolerant Gux1 peptide from A. cellulolyticus. The Gux1 exoglucanse is a member of the glycoside hydrolase family and comprises a caralytic domain GH48. carbohydrate binding domain type III, and a carbohydrate binding domain type III. The polypeptide is useful in the degradation of cellulose into biofuel, or pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial-scale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Gux1 cellulase CBD_III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
589 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18 DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN
                                                                                                                                                                                                                                                                        Gux1, thermal tolerant, exoglucanase, glycoside hydrolase, cellulose,
biofuel, detergent, pulp, paper processing, feed processing, textile,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 71; DB 6; Length 155/
Pred. No. 6e-64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Decker SR;

    A. cellulolyticus Gux1 protein CBD_III domain fragment.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79.8%; Scor.
100.0%; Pred. No. c.
... 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME,
                                                   Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant TB,
                                                                                                                                             ABP71658 standard; protein; 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 2; Page 7; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                                                                                                                                                       Acidothermus cellulolyticus.
                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2003-300494/29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                         cellulase; enzyme
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 153 AA;
                                                                                                                                                                                                                                                                                                                                                                       WO2003012095-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          domain fragment
                                                                                                                                                                                                            29-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                       13-FEB-2003
                                                                                                                                                                           ABP71658;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
```

Decker SR;

Himmel ME,

ö

Gaps . 0

Length 1121; 0; Indels

DB 6; Le 3.6e-63;

RESULT

a

PATPTÁDTYLO 87

a ò g

ò

RESULT 8

```
/note= "unspecified residue"
                                                          Location/Qualifiers
                                                                                                                                                                                                                                 8-JUL-2001; 2001WO-US023818.
                                                                                                                                                                                                                                                                      28-JUL-2001; 2001WO-US023818.
                   Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                    (MIDE ) MIDWEST RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-248177/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO2003012090-A2
                                                                                Misc-difference
                                                                                                                                                WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 89 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           03-JUN-2003
                                                                                                                                                                                         13-FEB-2003.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABP73015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               host
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
ABP73015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
  DX LL LL LL X S X K K K X C E X LL X S X X S X K K K X X C E X LL X S X K K K X X C E X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X LL X S X 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         designated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuels additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide
                                                                                                                                                                                Avicelase, AviIII; glycoside hydrolase; enzyme; cellulase; biofuel;
detergent; pulp processing; paper processing; feed processing; textile;
cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile; cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polynucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino acid sequence of AviIII carbohydrate binding domain type III.
                                                                                                                                                Amino acid sequence of AviIII carbohydrate binding domain type II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      present sequence is derived from a thermostable avicelase,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40.4%; Score 36; DB 6; L6
100.0%; Pred. No. 1.3e-28;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28 UNICSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Himmel ME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABP73017 standard; peptide; 89 AA
                        standard; peptide; 88 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 6; Page 8; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                         28-JUL-2001; 2001WO-US023818.
                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                      (first entry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 100.
nes 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MIDE ) MIDWEST RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2003-248177/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 88 AA;
                                                                                                                                                                                                                                                                                                                    WO2003012090-A2
                                                                                                        03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                 13-FEB-2003
                     ABP73018
                                                               ABP73018;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABP73017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

```
designated
                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is derived from a thermostable avicelase, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from generically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing AviIII polypeptide actions of the polypeptide. The AviIII polymuclectide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                             useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avicelase, AviIII; glycoside hydrolase, enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                             New thermostable AviIII peptide from Acidothermus cellulolyticus, usef
for degradation of cellulose or in generating anti-AviIII antibodies f
purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40.4%; Score 36; DB 6; Length 89; 100.0%; Pred. No. 1.3e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "unspecified residue encoded by N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sequence of the avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Pred. ...
   Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ź
Ding S, Adney WS, Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABP73015 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                         Claim 6; Page 8; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36; Conservative
```

ö

```
(DING/) DING S.
(ADNE/) ADNEY W S.
(VINZ/) VINZANT T B.
(HIMM/) HIMMEL M E.
      PARTICULAR SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present sequence represents a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed production of detergents, pulp and parer processing, food and feed useful in the degradation of cellulose, and in generating specific anti-aviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                     New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 957;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 6; Le
                                                                                                                                                                                                                                                                                               Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 36;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40.4%; Scor
100.0%; Pred
0; N
                                                                                                                                                                                                                                                                                               Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 2; Page 20; 44pp; English
                                                                                             28-JUL-2001; 2001WO-US023818.
                                                                                                                                                                28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                               (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                        2003-248177/24.
                                                                                                                                                                                                                                                                                               Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; ABZ77632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 957 AA;
                                    13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                           Ding S,
```

New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AvilII and for reducing cellulose in a starting material, e.g. municipal solid waste.

Claim 16; SEQ ID NO 1; 29pp; English

Himmel ME;

Vinzant TB,

Adney WS,

Ding S,

WPI; 2003-810853/76. N-PSDB; ADD22922.

```
ö
The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                      Length 957;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         896 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 931
                                                                                                                                                                                                                                                                                                                                                                                      DB 7; Le
1.2e-27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28 UNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                      40.4%; Score 36; DB 100.0%; Pred. No. 1.2
                                                                                                                                                                                                                                                                         cellulolyticus avicelase AviIII
                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 100.
Matches 36; Conservative
                                                                                                                                                                                                                                                                                                                               Sequence 957 AA;
```

Search completed: May 11, 2004, 12:26:52 Job time : 55 secs

o;

Gaps

ö

Indels

Mismatches

Conservative

36;

Matches

63

896 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 931

БЪ ઠ

28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA

enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.

Location/Qualifiers /label= Unknown /note= "Encoded by

Misc-difference

/note=

US2003108988-A1

12-JUN-2003

Acidothermus cellulolyticus

18-OCT-2002; 2002US-00155400.

28-JUL-2001; 2001US-00917376

Acidothermus cellulolyticus avicelase AviIII

(first entry)

15-JAN-2004

ADD22921;

ADD22921 standard; protein; 957 AA

RESULT 12

The

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 12:23:57; Search time 21 Seconds (without alignments) 407.669 Million cell updates/sec

Title: Perfect score: Sequence:

US-09-917-376-4 89 1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATFTADTYLQX 89

OLIGO Gapop 60.0 , Gapext 60.0 Scoring table:

283366 seqs, 96191526 residues

Searched:

10

Word size :

0 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description Query Score Match Length DB Result

No matches found

Search completed: May 11, 2004, 12:28:43 Job time : 21 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 12:11:27; Search time 15 Seconds (without alignments) 308.950 Million cell updates/sec

US-09-917-376-4 89 1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX 89 Title: Perfect score: Sequence:

Scoring table:

OLIGO Gapop 60.0 , Gapext 60.0

141681 segs, 52070155 residues Searched:

10 Word size :

0 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

SwissProt_42:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description Ouery Score Match Length DB No. Result

Search completed: May 11, 2004, 12:27:19 Job time: 16 secs

No matches found

```
May 11, 2004, 12:23:22 ; Search time 39 Seconds (without alignments) 720.029 Million cell updates/sec
                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX 89
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                  1017041 segs, 315518202 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Listing first 1000 summaries
                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                              OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
5: sp_inwan:*
6: sp_nammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_archeap: *
                                                                                                                                                                        US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPIREMBL 25:*
                                                                                                                                                                                                                                                                                                                                    10
                                                                                                                                                                          Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                  Word size :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database :
                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                       Run on:
```

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
```

SUMMARIES

H Query Score Match Length DB Ņ. Result

Description

No matches found

Search completed: May 11, 2004, 12:28:10 Job time : 39 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
OM protein - protein search, using sw model
```

May 11, 2004, 12:25:52; Search time 23 Seconds (without alignments) 199.770 Million cell updates/sec

Run on:

US-09-917-376-4 89 1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX Title: Perfect score: Sequence:

8

389414 segs, 51625971 residues OLIGO Gapop 60.0 , Gapext 60.0 Scoring table:

10 Word size : Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

1: /cgn2 6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:* Issued_Patents AA:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Score Match Length DB No. Result

Description

No matches found

Search completed: May 11, 2004, 12:29:19 Job time : 23 secs

```
(without alignments)
602.523 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. (cgn2_6/ptodata/2/pubpaa/USO7_PUBCOMB.pep:*
2. (cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO6_NEW_PUB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO8_NEW_PUB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
3. (cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
4. (cgn2_6/ptodata/2/pubpaa/USO8_PUBCOMB.pep:*
5. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
5. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO9_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO09_PUBCOMB.pep:*
6. (cgn2_6/ptodata/2/pubpaa/USO08_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            May 11, 2004, 12:28:17 ; Search time 41 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1140673 segs, 277566755 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Listing first 1000 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Published Applications AA:*
                                                                                                            OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                          OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                        US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10
                                                                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Word size
                                                                                                                                                                                                                                                                                                                                          Seguence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                     Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Appli													
-	'n	'n	4,	4,	4	'n	'n	'n	ý	ģ	'n	٦,	1,	4
Description	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Sequence	Seguence
ID	US-09-917-376-5	US-10-155-400-5	US-09-917-376-4	US-10-155-400-4	US-09-917-378-4	US-09-917-378-1	US-09-917-384-5	US-09-917-383-5	US-09-917-384-6	US-09-917-383-6	US-09-917-384-1	US-09-917-383-1	US-09-917-376-1	US-10-155-400-1
	12	14	12	14	10	10	10	10	10	10	10	10	12	14
% Query Match Length DB	88	88	89	8	154	762	150	150	1043	1043	1228	1228	957	957
% Query Match	98.9	98.9	98.9	98.9	98.9	98.9	93.3	93.3	93.3	93.3	93.3	93.3	40.4	40.4
Score	88	88	88	88	88	88	83	83	83	83	83	83	36	36
Result No.	1	7	m	4	S	9	7	œ	σ	10	11	12	13	14

ALIGNMENTS

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
      Sequence 5, Application US/09917376

Publication No. US20040038334A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: VINZANT, TODD B.

APPLICANT: VINZANT, TODD B.

APPLICANT: HIMMEL, MICHAEL E.

APPLICANT: HIMMEL, MICHAEL E.

APPLICANT: HIMMEL, MICHAEL E.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

CURRENT APPLICANT: 2001-07-28

NUMBER OF SEQ ID NOS: 7

SEQ ID NO 5

LENGTH: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 5, Application US/10155400

Sequence 5, Application US/10155400

Publication No. US2033108988A1

GENERAL INFORMATION:

APPLICANT: DING, SHIL-YOU

APPLICANT: WILLIAM S.

APPLICANT: VINZANT, TODD B.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

CURRENT FALING DATE: 2002-10-22

NUMBER OF SEQ ID NOS: 7

SEQ ID NO S.

LENGTH: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98.9%; Score 88; DB 12; ] 100.0%; Pred. No. 1.7e-78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98.9%; Score 88; DB 14; 1 100.0%; Pred. No. 1.7e-78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION: Carbohydrate binding domain
                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION: Carbohydrate binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Prec. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          l Similarity 100.0%; Pred. No. 1.7
88; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 88; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 88; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
US-10-155-400-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-10-155-400-5
US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                   FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

Length 89;

g

à

```
98.9%; Score 88; DB 14; I
100.0%; Pred. No. 1.7e-78;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 100.0
                Local Similarity 100.
nes 88; Conservative
                                                                                                                                                                                                                                           US-09-917-378-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-378-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 1
LENGTH: 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-378-1
 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE
                   Best Loc
Matches
                                                                                                     셤
                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <del>a</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, TODD B.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: NREL 01-36A
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PALENTIN Ver. 2.1
SEQ ID NO 4
LENGTH: 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match 98.9%; Score 88; DB 12; I Best Local Similarity 100.0%; Pred. No. 1.7e-78; Matches 88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: Any amino acid
OTHER INFORMATION: Carbohydrate binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-4
                   61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
NAME/KEY: MOD RES
LOCATION: (89)
OTHER INFORMATION: Any amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 4, Application US/10155400; Publication No. US20030108988A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: MOD RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-155-400-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
```

g

8

g

ð

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGCVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4, Application US/09917378

Publication No. US20030119093A1

GENERAL INFORMATION:

APPLICANT: DING, SHIL-YOU

APPLICANT: DING, SHIL-YOU

APPLICANT: VINZANT, TODD B.

APPLICANT: WILLIAM S.

APPLICANT: HIRMEL, MICHAEL E.

APPLICANT: HIRMEL, MICHAEL E.

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 4097.7US01

CURRENT APPLICATION NUMBER: US/09/917,378

CURRENT FILING DATE: 2001-07-28

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/08917378
| Publication No. US20030119093A1
| GENERAL INFORMATION:
| APPLICANT: DING, SHI-YOU
| APPLICANT: DING, SHI-YOU
| APPLICANT: DING, STEPHEN S.
| APPLICANT: DECKER, STEPHEN R.
| APPLICANT: HIMMEL, MICHAEL E.
| APPLICANT: HIMMEL, MICHAEL E.
| APPLICANT: HIMMEL, MICHAEL E.
| APPLICANT: HIMMEL, MICHAEL E.
| APPLICANT: HIMMEL, MICHAEL E.
| TITLE OF INVENTION: CELLULOLYTICUS
| TITLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF INVENTION: CELLULOLYTICUS
| TILLE OF I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate OTHER INFORMATION: binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.9%; Score 88; DB 10; I 100.0%; Pred. No. 2.7e-78; cive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
```

```
TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-383-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-384-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                           셤
                                                                                     ö
                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                          1 VSGGVKVQYKNIDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 5, Application US/09917383;
Publication No. US20030104522A1
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40770.62801
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: VINZAN', WILLIAM S.
APPLICANT: VINZAN', TODIB
APPLICANT: PROBLEMS.
APPLICANT: PROBLEMS.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170.6US01
CURRENT APPLICATION NUMBER: US/09/917,384
CURRENT APPLICATION NUMBER: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARR: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; OTHER INFORMATION: Description of Artificial Sequence: Segment of; OTHER INFORMATION: GuxA
US-09-917-384-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
0
                                                                                     ö
                                         98.9%; Score 88; DB 10; Length 762; 100.0%; Pred. No. 1.1e-77; tive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 93.3%; Score 83; DB 10; Length 150; Best Local Similarity 100.0%; Pred. No. 2.1e-73; Matches 83; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 5, Application US/09917384; Publication No. US20030096342A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Artificial Sequence
                                                                                     88; Conservative
                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 5
LENGTH: 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                             US-09-917-384-5
  US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                        유
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                          g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                                                                                                                                                                                                         65
                                                                                                                                                                                                                                                                                  6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAWG
                                                                                                                                                                                                                                                        6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                 Gaps
PEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Segment of
CTHER INFORMATION: GuxA
US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ) OTHER INFORMATION: Description of Artificial Sequence: Segment of CTHER INFORMATION: GuxA
US-09-917-384-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 6, Application US/09917383
Publication No. US2003010452241
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS TITLE OF INVENTION: CELLULLYTICUS
FILE REPERENCE: 40170-61801
CURRENT APPLICATION NUMBER: US/09/917,383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: DING, SH.-YOU
APPLICANT: DING, SH.-YOU
APPLICANT: DING, SH.-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINCANT, TOOD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERNAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REPERENCE: 40170.6US01
CURRENT APPLICATION NUMBER: US/09/917,384
CURRENT APPLICATION NUMBER: US/09/917,384
SURRENT PILLING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                 ;;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1043;
                                                                                                                                          Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          93.3%; Score 83; DB 10; Length 10
100.0%; Pred. No. 1.1e-72;
tive 0; Mismatches 0; Indels
                                                                                                                                    Query Match 93.3%; Score 83; DB 10; Length 15
Best Local Similarity 100.0%; Pred. No. 2.1e-73;
Matches 83; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             542 CGNIRASFGSVNPATPTADTYLQ 564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                     66 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                         66 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 6, Application US/09917384 Publication No. US20030096342A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 93.3
Best Local Similarity 100.
Matches 83; Conservative
```

à 임 ò g

```
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     649 CGNIRASFGSVNPATPTADTYLQ 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: MOD_RES

LOCATION: (957)

OTHER INFORMATION: Any amino acid

US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/10155400 ; Publication No. US20030108988A1 ; GENERAL INFORMATION:
                                                                                                                                                                                                                           TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: DING, SHI-YOU APPLICANT: ADNEY, WILLIAM S. APPLICANT: VINZANT, TODD B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                         US-09-917-383-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                       482 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDMAAMG 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          589 KVQYKNNDSAPGDNOIKPGLOLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNCDWAAMG 648
                                                                                                                                                                                                                                                                                                                                                              65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65
                                                                                                                                                                                                                                                                                                                                                           6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                             OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                  .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERWAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION:
FILE REFERENCE: 40170.6US01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                  93.3%; Score 83; DB 10; Length 1043; 100.0%; Pred. No. 1.1e-72; tive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 93.3%; Score 83; DB 10; Length 1228; Best Local Similarity 100.0%; Pred. No. 1.3e-72; Matches 83; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION NUMBER: US/09/917,384
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/09917383; Publication No. US20030104522A1; GENERAL INFORMATION:
A PAPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1, Application US/09917384; Publication No. US20030096342A1; GENERAL INFORMATION:
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PALENLIN Ver. 2.1
SEQ ID NO 6
LENGTH: 1043
                                                                                                             TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                           Best Local Similarity 100. Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH: 1228
                                                                                                                                                                                                                           US-09-917-383-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-383-1
                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
                                                                                                                                                           FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
```

```
589 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/08917376
Publication No. US20040038334A1
GENERAL INFORMATION:
APPLICANT: DING, SHILYION
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: THEMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.4US01
                                                                                                                                                                                                                                                                                                                                                                                    OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ,
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .
0
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170. 120/09/917,383
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 1
LENGTH: 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1228;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 957;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 93.3%; Score 83; DB 10; L
Best Local Similarity 100.0%; Pred. No. 1.3e-72;
Matches 83; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        896 VNIGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.3e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tch 40.4%; Score 36; DB 12; al Similarity 100.0%; Pred. No. 1.3e-26, 36; Conservative 0; Mismatches 0;
```

ö

g Š g

```
## APPLICANT: HIMMEL, MICHAEL E.
## TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
## TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
## CURRENT APPLICATION UNMERS: US/10/155,400
## CURRENT APPLICATION NUMBER: US/10/155,400
## CURRENT APPLICATION NUMBER: US/10/155,400
## CURRENT APPLICATION NUMBER: US/10-22
## CURRENT APPLICATION NUMBER: US/10-10-22
## CORRIGION OF 1
## CORTORS OF SECTION OF 1
## CORTORS OF INTERPRETATION OF 1
## CORTORS OF INTERPRETATION OF 1
## CORTORS OF INTERPRETATION OF 1
## CORTORS OF INTERPRETATION OF 1
## CORSENTATION OF 1
## CORTORS OF INTERPRETATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION OF 1
## CORSENTATION
```

Appli Appli Appli Appli

Sequence s Sequence Sequence Sequence Sequence S Sequence Sequence

Appli Appli Appli Appli Appli

4444400001

Appli Appli Appli Appli Appli

Sequence 1 Sequence 4 Sequence 1

Sequence 1,

Sequence

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
US-10-155-400-5
; Sequence 5, Application US/10155400
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: DINGX, WILLIAM S.
; APPLICANT: VINZANT: TODB B.
; APPLICANT: HIMMEL, MICHEL B.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: NREL 01-362
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PATCHILIN Ver. 2.1
; SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, MICHAEL E.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: CELULOLYTICUS
TITLE OF INVENTION: CELULOLYTICUS
TITLE OF INVENTION: CELULOLYTICUS
TITLE OF INVENTION: 0.001-07-28
CURRENT APPLICAND NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98.9%; Score 88; DB 24;
llarity 100.0%; Pred. No. 8e-81;
Conservative 0; Mismatches 0
                                    US-09-917-376-4
US-10-155-400-4
US-09-917-378-4
US-09-917-378-1
US-09-917-384-5
US-09-917-384-6
US-09-917-384-6
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384A-1
US-09-917-384A-1
US-09-917-384A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; OTHER INFORMATION: Carbohydrate binding domain US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5; Sequence 5, Application US/09917376; GENERAL INFORMATION:
  88
89
89
89
154
1062
11043
11043
11228
11228
11238
11238
11238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 88; Conserv
FEATURE:
                                             645969876543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                  May 11, 2004, 12:26:57 ; Search time 182 Seconds (without alignments) 477.301 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pending Patents AA Main:*

1: (ggn2_6/ptodata/2/paa/USG6_COMB.pep:*

2: (cgn2_6/ptodata/2/paa/USG6_COMB.pep:*

3: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

5: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

6: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

6: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

9: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

11: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

12: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

13: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

14: (cgn2_6/ptodata/2/paa/USG8_COMB.pep:*

15: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

16: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

16: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

17: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

18: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

19: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG9_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

11: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

12: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

13: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

14: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

15: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

16: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

17: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

18: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

19: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (cgn2_6/ptodata/2/paa/USG0_COMB.pep:*

10: (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
                    GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                               6019581 segs, 976053577 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Post-processing: Listing first 1000 summaries
                                                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                                                                                OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length: 0
length: 2000000000
                                                                                                                                                                                                                                      US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Score Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                            2
```

Minimum DB seq Maximum DB seq

Word size

Searched:

Database :

Title: Perfect score:

Sequence:

OM protein -

Run on:

Scoring table:

ö

ö

Length 88; 0; Indels 9 9

Description

No. Result

```
PatentIn Ver. 2.1
SOFTWARE: Par
SEQ ID NO 4
LENGTH: 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
US-09-917-378-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                 Query Match
                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.4USO1
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT APPLICATION NUMBER: US/09/917,376
SOUTHER OF SEQ ID NOS: 7
SOFTWARE: PALENTIN VET. 2.1
SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHEL E.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEQ ID NOS: 7
                                                                                                                                                          98.9%; Score 88; DB 27; Length 88; 100.0%; Pred. No. 8e-81; ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98.9%; Score 88; DB 24; Length 89; 100.0%; Pred. No. 8.1e-81; Live 0; Mismatches 0; Indels
                                                                                      ; OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OTHER INFORMATION: Any amino acid OTHER INFORMATION: Carbohydrate binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                           61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                     61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                         TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Sequence 4, Application US/10155400
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 4, Application US/09917376 GENERAL INFORMATION:
                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 100.0
Matches 88; Conservative
                                                                                                                                                     Query Match
Best Local Similarity
Matches 88, Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: MOD RES
LOCATION: (89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
US-10-155-400-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT
                                                                      FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
  LENGTH:
                                                                                                                                                                                                                                                                                          임
                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                     à
```

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                          1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate; OTHER INFORMATION: binding domain US-09-917-378-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: DINGY, WILLAM S.
APPLICANT: DINGY, WILLAM S.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.70501
CURRENT APPLICATION NUMBER: US/09/917,378
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 154;
                                                                                                                                                                                                                                                  Length 89;
                                                                                                                                                                                                                                                  DB 27; I
8.1e-81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98.9%; Score 88; DB 24; I
llarity 100.0%; Pred. No. 1.3e-80;
Conservative 0; Mismatches 0;
                                                                                                                                                                 , OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-4
                                                                                                                                                                                                                                                ch 98.9%; Score 88; DB 3 Similarity 100.0%; Pred. No. 8.1 88; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
ORGANISM: Acidothermus cellulolyticus FEATURE:
NAME/KEY: MOD RES
LOCATION: (89)
OTHER INFORMATION: Any amino acid
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 4, Application US/09917378; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/09917378; GENERAL INFORMATION:
APPLICANT: DING, SH1-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN B.
APPLICANT: HIMMEL, MICHAEL B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
hes 88; Conserv
                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 88; Conserv
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                           455 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: HIMMEL, TODD B.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: TERNAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: GELJOLO
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT APPLICATION NUMBER: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO S
LENGTH: 150
TYPE: PRT
OGRANISM: Artificial Sequence
TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
TITLE OP INVENTION: CELLULOLYTICUS
FILE REPERENCE: 40197.7010
CURRENT APPLICATION NUMBER: US/09/917,378
CURRENT PILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver: 2.1
SEQ ID NO 1
LENGTH: 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER INFORMATION: Description of Artificial Sequence: Segment OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93.3%; Score 83; DB 24; Length 150; 100.0%; Pred. No. 1.6e-75;
                                                                                                                                                                                                                                                                                                                DB 24; Length 762;
5.8e-80;
                                                                                                                                                                                                                                                                                                                Query Match 98.9%; Score 88; DB Best Local Similarity 100.0%; Pred. No. 5.8 Matches 88; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                     ; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5, Application US/09917384
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL, MICHAEL E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-383-5; Sequence 5, Application US/09917383; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 100. Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-384-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6 KVQYKNNDSAPGDNQ1KPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 6, Application US/09917383
Sequence 6, Application US/09917383
GENERAL INFRMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HUMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULLOLYTICUS
FILE REFERENCE: 40170.6US01
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT APPLICATION NUMBER: US/09/917,383
SOFTWARE: PATENTING DATE: 2001.07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTING DATE: 201.07-28
LENGTHEN: 1043
                                                                                                                                                                                                                                                                                                               ; OTHER INFORMATION: Description of Artificial Sequence: Segment of ; OTHER INFORMATION: GuxA
US-09-917-384-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
COTHER INFORMATION: Description of Artificial Sequence: Segment of
COTHER INFORMATION: GuxA
US-09-917-383-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS TITLE OF INVENTION: CELLULOLYTICUS FILE REFERENCE: 40170.6US01
CURRENT APPLICATION NUMBER: US/09/917,384
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1043;
                                                                                                                                                                                                                                                                                                                                                                                                                      Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93.3%; Score 83; DB 24; Length 10
100.0%; Pred. No. 9.4e-75;
ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                 93.3%; Score 83; DB 24; Length 15
100.0%; Pred. No. 1.6e-75;
tive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-384-6
Sequence 6, Application US/09917384
GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
HAPPLICANT: ADNEY, WILLIAM S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                         TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 100.
Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
```

ö

```
US-09-917-384A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
US-09-917-384A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 482 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6 KVOYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
## APPLICANT: VINZANT, TODD B.
## APPLICANT: DECKER, STEPHEN R.
## APPLICANT: HIMMEL, MICHAEL R.
## TITLE OF INVENTION: CELLULOLYTICUS
## TITLE OF INVENTION: CELLULOLYTICUS
## FILE REFERENCE: 40170.60801
## CURRENT APPLICATION NUMBER: US/09/917,384
## CURRENT PILING DATE: 2001-07-28
## NUMBER OF SEQ ID NOS: 14
## SEQ ID NO 6
## SEQ ID NO 6
## LENGTH: 1043
## TYPE: PRT
## ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TOOD B.
APPLICANT: VINZANT, TOOD B.
APPLICANT: HIWMEL, MICHAEL E.
TITLE OF INVENTION: THERNAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170.6US01
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOUTHWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1228
                                                                                                                                                                                                                                                                                                                                                            ; OTHER INFORMATION: Description of Artificial Sequence: Segment of; OTHER INFORMATION: GuxA
US-09-917-384-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 93.3%; Score 83; DB 24; Length 1043; Best Local Similarity 100.0%; Pred. No. 9.4e-75; Matches 83; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                93.3%; Score 83; DB 24; Length 1228; 100.0%; Pred. No. 1.1e-74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CGNIRASFGSVNPATPTADTYLQ 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66 CGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09917383 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Sequence 1, Application US/09917384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 100.0
Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-383-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            589 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17 DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6 KVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: MCCARTER, SUZANNE
APPLICANT: MCCARTER, SUZANNE
APPLICANT: HIMMEL, MCHAEL E.
APPLICANT: HIMMEL, MCHAEL E.
APPLICANT: VINZANT, TODD B.
TITLE OF INVENTION: THERMAL FOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: NREL 01-38
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PATENTIN VOT: 2.1
SOFTWARE: PATENTIN VOT: 2.1
SEQ ID NO 4
LENGTH: 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18 DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
APPLICANT: DING, SHI-YOU
APPLICANT: DING, MILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZAN, TODD B.
APPLICANT: DECKEN, STEPHEN R.
APPLICANT: DECKEN, STEPHEN R.
APPLICANT: DECKEN, STEPHEN R.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS;
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170.6US01
CURRENT PRILITY DATE: 201-07-28
NUMBER OF SEQ ID NOS: 14
SOUTHARE: PALENTING DATE: 201-07-28
SOUTHARE: PALENTING DATE: 201-07-28
SOUTHARE: PALENTING DATE: 201-07-28
SOUTHARE: PALENTING DATE: 201-07-28
SOUTHARE: PALENTING DATE: 201-07-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; OTHER INFORMATION: Description of Artificial Sequence: Segment of CTHER INFORMATION: GuxA
US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1228;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.3%; Score 83; DB 24; I
100.0%; Pred. No. 1.1e-74;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79.8%; Score 71; DB 24; 1
100.0%; Pred. No. 2.5e-63;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Acidothermus cellulolyticus US-09-917-384A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66 CGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 4, Application US/09917384A GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ilarity 100.0%; P
Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 93.3
Best Local Similarity 100.
Matches 83; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PATPTADTYLO 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78 PATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 71; Conserv
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 51 DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18 DNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGCGNIRASFGSVN 77
                                                                                                    APPLICANT: DING, SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJANG SHIJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: DING; SHI-YOU
APPLICANT: DING; SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT: TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REPERENCE: 40197.41801
CURRENT FILING DATE: 2001-07-28
CURRENT FILING DATE: 2001-07-28
SOFTWARE: PATENTIN VOS: 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40.4%; Score 36; DB 24; Length 957; 100.0%; Pred. No. 4.9e-27; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            896 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
79.8%; Score 71; DB 24; I
Best Local Similarity 100.0%; Pred. No. 1.6e-62;
Matches 71; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: PRT ORGANISM: Acidothermus cellulolyticus
Application US/09917384A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:

NAME/KEY: MOD RES

LOCATION: (957)

OTHER INFORMATION: Any amino acid
US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1, Application US/09917376 GENERAL INFORMATION:
                                                                                          APPLICANT: ADNEY, WILLIAM S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 100.
Matches 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111 PATPTADTYLO 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78 PATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-384A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 1
LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠ
```

Sequence 1, Application US/10155400; GENERAL INFORMATION: , APPLICANT: DING, SHI-YOU

US-10-155-400-1

```
APPLICANT: WILLIAM S.

APPLICANT: HINZANT, TODD B.

TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: NREL OI - 3 6A

CURRENT APPLICATION NUMBER: US/10/155,400

CURRENT FILING DATE: 2002-10-22

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PATENTIN Ver. 2.1

SEQ ID NO 1

LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 957;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     896 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 931
                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 36; DB 27;
Pred. No. 4.9e-27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28 VNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                  40.4%; Scor.
100.0%; Pred. No. 2..
0; Mismatches
                                                                                                                                                                                                                                                                                                      TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: May 11, 2004, 12:32:34
Job time : 183 secs
                                                                                                                                                                                                                                                                                                                                         FEATURE:

NAME/KEY: MOD RES
LOCATION: (957)
COTHER INFORMATION: Any amino acid
US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
OM protein - protein search, using sw model
```

May 11, 2004, 12:27:22 ; Search time 18 Seconds (without alignments) 169.758 Million cell updates/sec 1 VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX 89 US-09-917-376-4 89 Title: Perfect score: Sequence: Scoring table: Run on:

234425 segs, 34333021 residues OLIGO Gapop 60.0 , Gapext 60.0 Searched:

Total number of hits satisfying chosen parameters: 10 Word size :

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query Score Match Length DB М О Result

Description

No matches found

Search completed: May 11, 2004, 12:33:04 Job time : 18 secs

```
PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                     AX700050 Sequence
AX700036 Sequence
AX700025 Sequence
AX700058 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Acidothermus cellulolýticus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Frankineae; Acidothermaceae; Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ding, S.Y., Adney, W.S., Vinzant, T.B. and Himmel, M.E.
Thermal tolerant mannanase from Acidothermus cellulolyticus
Patent: WO 03012110.4 2 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. .2289
/organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2289
88
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AX700050 2289 bp Sequence 2 from Patent W003012110.
                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                        AX700050
AX700036
AX700025
AX700058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89) x AX700050 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Acidothermus cellulolyticus
                                                                                                                                                      em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*
                            em_htg_inv:*
em_htg_other:*
em_htg_mus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AX700050.1 GI:29536020
                                           em htg_other:
em htg_mus:*
em htg_pln:*
em htg_rod:*
em htg_mam:*
em htg_vrt:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.02e-83
88.00
100.00%
100.00%
98.88%
               htg hum:*
                                                                                                                                                                                                                                                                                                            Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                            99
em_vi:*
em_htg_h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                        98.9
93.3
79.8
40.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AX700050
39:
                                                                                                                                                                      40:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                                                                                                                                                          88
83
71
36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
AX700050
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -MODEL=frame+ pan.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/USO9917376/runat_11052004_114646_28658/app_query.fasta_
-Q=/cgn2_1/USPTO_spool/USO9917376/runat_11052004_114646_28658/app_query.fasta_
-DB=censmbl -QFWT=fastap -SUFFTX=clip2n.rge -WINMATCH=0.1.cLOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=cligo -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=quality -THR MIN=30 -ALIGM=50 -MODE=LOCAL -OUTFWT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRE-USRG-STAFF 6 GGA 1 1.371 @runat 11052004 114646_28658 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELEXT=7
                                                                                                      May 12, 2004, 14:59:07; Search time 2808 Seconds (without alignments) 1373.764 Million cell updates/sec
                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
               GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                       OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                            3470272 seqs, 21671516995 residues
                                                                                                                                                                                                                                                                                                                                                                                       Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Listing first 1000 summaries
                                                                                                                                                                                                                                 OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                 US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gb ba: *
gb htg: *
gb ln: *
gb ow: *
gb by: *
gb pl: *
gb pl: *
gb re: *
gb v: *
gb v: *
gb v: *
gb v: *
gb v: *
gb w: *
gb w: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_pat: *
em_ph: *
em_ph: *
em_pl: *
em_ro: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_in:*
em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_om:*
em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em ov: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GenEmbl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10:
112:
14:
                                                                                                                                                                                                                                                                                                                                                            30
                                                                                                                                                                    Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                                                                                                                                                                                                                                                                            Word size:
                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                          Run on:
```

```
Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                               TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                            REFERENCE
                                         AUTHORS
                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AX700058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                1542
                                                        1543 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                    03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1885 Trcacccecarecrecercercacacacreeraracaacrereacreecececeared
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1765 AAGGIGCAGIACAAGAACAAIGATICGGCGCCGGGIGATAACCAGATCAAACCGGGTCTC 1824
                                       80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                            Ding, S.Y., Adney, W.S., Vinzant, T.B., Himmel, M.E. and Decker, S.R. Thermal tolerant cellulase from Acidothermus cellulolyticus Patent: WO 03012109-A 2 13-FEB-2003; Midwest Research Institute (US)
1483 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LysValGlnTyrLysAsnAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26 GlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyrTrp
                                       61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr
                                                                                                                                                                    PAT
                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                 3687
83
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                           organism="synthetic construct"
                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
                                                                                                                                                                                                                                                                                                                                                                     /mol_type="unassigned DNA"
/db_xref="taxon:32630"
                                                                                                                                                                 AX700036 3687 bp
Sequence 2 from Patent WO03012109.
AX700036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AX700025 3365 bp
Sequence 2 from Patent WO03012095.
AX700025.1 GI:29536018
                                                                                                                                                                                                                                                                                                                                                                                               /note="Segment of GuxA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                              1603 CCGACGCCGACACCTACCTGCAG 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4 (1-89) x AX700036 (1-3687)
                                                                                        81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus
Acidothermus cellulolyticus
                                                                                                                                                                                                      AX700036.1 GI:29536019
                                                                                                                                                                                                                                                                                                                                                                                                                                              3.23e-78
83.00
100.00%
100.00%
93.26%
                                                                                                                                                                                                                             synthetic construct
synthetic construct
artificial sequences.
                                                                                                                                                                                                                                                                                                                                               . .3687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ||||||||||
TACCTGCAG 2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                                       AX700036
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                               AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
AX700025
LOCUS
                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                          g
                                                                                        ð
                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
```

```
PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       151 GACAACCAGATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGTCGTGGATTTG 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               211 TCGACGGTGACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTAC 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    271 AACTGTGACTGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAAC 330
                                                         Adney, W.S., Ding, S.Y., Vinzant, T.B., Himmel, M.E., Decker, S.R. and Lantz Mccarter, S.
Lantz Mccarter, S.
Thermal tolerant exoglucanase from Acidothermus cellulolyticus Patent: Woll 2012 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18 AspAsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeu 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 38 SerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyr 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        58 AsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Sacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ding, S.Y., Adney, W.S., Vinzant, T.B. and Himmel, M.E.
Thermal tolerant avicelase from Acidothermus cellulolyticus
Patent: WO 033012090-A 2 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                      /organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2869
36
0
0
0
                                                                                                                                                                                                                                                                                                                                                          3365
71
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Frankineae, Acidothermaceae, Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                       Frankineae; Acidothermaceae; Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          331 CCGGCGACGCCGACGCGGACACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  78 ProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AX700058 . 2869 bp
Sequence 2 from Patent WO03012090.
AX700058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-4 (1-89) x AX700025 (1-3365)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-4 (1-89) x AX700058 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Acidothermus cellulolyticus
Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AX700058.1 GI:29536021
                                                                                                                                                                                                                                                                                                                                                       1.82e-65
71.00
100.00%
100.00%
79.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.09e-28
36.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .. .2869
```

```
상 a · 상 a
```

Search completed: May 12, 2004, 16:05:36 Job time : 2810 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
OM protein - nucleic search, using frame_plus_p2n model
```

; Search time 322 Seconds May 12, 2004, 11:39:06 on:

Run

(without alignments)
1174.191 Million cell updates/sec

US-09-917-376-4

VSGGVKVQYKNNDSAPGDNQ.....RASFGSVNPATPTADTYLQX Perfect score: Sequence:

83

60.0 60.0 7.0 Xgapext Ygapext Fgapext 60.0 X 60.0 X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y X 60.0 Y Kgapop Fgapop (gapop OLIGO Scoring table:

3373863 segs, 2124099041 residues Searched:

Delext

Delop

30 Word size: Total number of hits satisfying chosen parameters:

ß

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Listing first 1000 summaries

Command line parameters:

-MODEL=frame+_p2n.model - DEV=xlh
-Q=/Cgn2_1/USPTO spool/US09917376/runat_11052004_114645_28650/app_query.fasta_1.263
-Q=/Cgn2_1/USPTO spool/US09917376/runat_11052004_114645_28650/app_query.fasta_1.263
-DB=MS Geneseq 295and4 -QPMT=fastap - SUFFIX=olip2n.rng -MINMATCH=0.1 - LOOPCL=0
-LLOOPEXT=0 -UMITS=bits - START=1 - END=-1 - MATRIX=oligo - TRANS=human40.cdi
-LIST=1000 -DOCALIGN=200 - THR_SCORE=quality - THR_MIN=30 - ALIGN=50 - MODE=LOCAL
-OUTFMT=pto -NORM=xt - HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09917376_GCN 1_470_GCN = MAYLEN=1052004_114645_28650 - NCPU=6 -ICPU=3
-NO MMAP - LARGEQUERY - NEG_SCORES=0 - WAIT - DSPBLOCK=100 -LONGIAG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 - THREADS=1 - XGAPOP=60 - XGAPEXT=60 - FGAPOP=6
-FGAPEXT=7 - YGAPOP=60 - YGAPEXT=60 - DELEXT=7

Geneseq 29Jan04:* Database

geneseqn2001bs:* geneseqn2001as:* geneseqn1980s:* geneseqn1990s:* geneseqn2000s:*

geneseqn2003bs:* geneseqn2003as:* geneseqn2003cs: geneseqn2004s:

geneseqn2002s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description		Abz77634 Nucleotid	Abz76162 A. cellul	Abz77632 Nucleotid	Add22922 Acidother
ΩI	ABZ77633	ABZ77634	ABZ76162	ABZ77632	ADD22922
	7	7	7	7	σ
Query Match Length DB		3687	3365	2869	2869
Query	98.9	93.3	79.8	40.4	40.4
Score	1 88 98.9	83	71	36	36
Result No.	1	7	٣	4	S.

ALIGNMENTS

40

```
The present sequence encodes ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1363 GTGTCGGGTGGGGTGAAGGTGCAGTACAAGAACAATGATTCGGCGCGGGGGTGATAACCAG 1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1423 ATCAAACCGGGFCTCCCAGFTGGFGAATACGGGGFTCGFTCGFCGGFGGAFTTGFCGACGGFG 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                      mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
feed; paper pulp; biofuel; mannase; gene; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463 A; 700 C; 688 G; 438 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2289
88
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                          Nucleotide sequence of the ManA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                x ABZ77633 (1-2289)
                                                                                                                                                                                                                                  Location/Qualifiers
1. 2289
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 1; Page 23; 46pp; English.
                                                                                                                                                                                                                                                                                       /product= "ManA"
                             ABZ77633 standard; DNA; 2289 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                      Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.32e-77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88.00
100.00%
100.00%
98.88%
                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-248182/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-4 (1-89)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-PSDB; ABP73022.
                                                                                                                                                                                                                                                                                                                      WO2003012110-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                            03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                      13-FEB-2003
                                                             ABZ77633;
                                                                                                                                                           ManA;
RESULT
                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

Indels:

```
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                  RESULT 3
                                                                  ò
                                                                                          요
                                                                                                                         ò
                                                                                                                                                   요
                                                                                                                                                                             à
                                                                                                                                                                                                      g
                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                    g
          1543 TGGGCGGCGATGGGGTGTGGGGAATATCCGCGCCTCGTTCGGCTCGTGGGGAACCGGCGACG 1602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence encodes a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                  80
                                                                                                                                                                                                                                                                                                                    GuxA; cellulase; glycoside hydrolase; enzyme; cellulose; sugar; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                  TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 3687 BP; 725 A; 1173 C; 1171 G; 618 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Decker SR;
                                                                                                                                                                                                                                                                                         Nucleotide sequence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Himmel ME,
                                                                                                                                      1603 CCGACGCGGACACCTACCTGCAG 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example 1; Page 23-24; 47pp; English.
                                                                                                                       ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1. .3687
/*tag= a
/product= "GuxA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Vinzant TB,
                                                                                                                                                                                                         BP
                                                                                                                                                                                                        ABZ77634 standard; DNA; 3687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                           Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                             03-JUN-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2003-239526/23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P-PSDB; ABP73029.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO2003012109-A1
                                                                  61
                                                                                                                       81
                                                                                                                                                                                                                                   ABZ77634;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding
                                                                                                                                                                                           ABZ77634
                                                                                                                                                                             RESULT
                                    셤
                                                                                         g
                                                                                                                                              g
                                                                                                                                                                                                         ò
                                                                                                                      ò
           8
```

3687 83 0

Length:
Matches:
Conservative:
Mismatches:

3.23e-72 83.00 100.00% 100.00%

Percent Similarity: Best Local Similarity:

Alignment Scores:

```
1885 TICACCCGGGATGGTGGTCGTCGACACTGGTGTACAACTGTACTGGCGGGGATGGGG 1944
                                                                                             1765 AAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAGATCAAACCGGGTCTC 1824
                                                                                                                                                        1825 CACTTGGTGAATACCGGGTCGTCGTCGGTGGATTGTCGACGGTGACGGTGCGGTACTGG 1884
                                                                                                                                                                                                                                                                                  1945 TGTGGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACGCCGACGGCGGAAACC
                                                                                                                                        45
                                                                                                                                                                                                    65
                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                                                               CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr
                                                                           LysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu
                                                                                                                                       GlnLeuValAsnThrGlySerSerValAspLeuSerThrValArgTyrTrp
                                                                                                                                                                                                     PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gux1; thermal tolerant; exoglucanase; glycoside hydrolase; cellulose; biofuel; detergent; pulp; paper processing; feed processing; textile; cellulase; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a thermal tolerant Gux1 peptide from A. cellulolyticus. The Gux1 exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /fransl_except= (pos: 682. .683, aa: pro)
/note= "this codon has an apparent one nucleotide
basepair deletion which alters the reading frame"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Decker
 00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A. cellulolyticus Gux1 protein encoding DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Page 22-23; 44pp; English.
                                             US-09-917-376-4 (1-89) x ABZ77634 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
1. .3365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      'product= "Gux1"
                                                                                                                                                                                                                                                                                                                                                                                                                      ABZ76162 standard; DNA; 3365 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         *tag= a
93.26%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                   TACCTGCAG 2013
                                                                                                                                                                                                                                                                                                                              TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2003-300494/29.
P-PSDB; ABP71656.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO2003012095-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                           2005
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABZ76162;
                                                                             9
                                                                                                                                        56
                                                                                                                                                                                                    46
                                                                                                                                                                                                                                                                 99
                                                                                                                                                                                                                                                                                                                              86
```

Himmel ME;

Vinzant TB,

```
Adney WS,
                                                                                                                                                                                                                                                                         Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADD22922;
  ŝ
                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                    Pred. No.:
  Ding
                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                              Score:
  ò
                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                  ð
binding domain type III, and a carbohydrate binding domain type II. The polypeptide is useful in the degradation of callulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial scale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Guxl cellulase encoding
                                                                                                                                                                                                                                                                                    151 GACAACCAGATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGTGGATTTG 210
                                                                                                                                                                                                                                  AspAsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeu 37
                                                                                                                                                                                                                                                                         57
                                                                                                                                                                                                                                                                                                               AsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Avicelase, AviIII; glycoside hydrolase, enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile; ceilulose; gene; ss.
                                                                                                                                                                                                                                                                        38 SerThrValThrValArgTyrTrpPheThrArgAspGlyGerSerThrLeuValTyr
                                                                                                             Sequence 3365 BP; 669 A; 1117 C; 1041 G; 538 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1. .2869
/*tag= a
/partial
/product= "AviIII"
/transl_except= (pos:2869,aa:Xaa)
/note= "Xaa is an unspecified residue"
                                                                                                                                          3365
71
0
0
0
                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                    CCGCCGACGCCGACGCGGACACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                       ProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nucleotide sequence of the avicelase AviIII.
                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                               US-09-917-376-4 (1-89) x ABZ76162 (1-3365)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                  BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023818.
                                                                                                                                                                                                                                                                                                                                                                                                                  ABZ77632 standard; DNA; 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023818
                                                                                                                                          2.29e-60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cellulolyticus.
                                                                                                                                                  71.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                79.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MIDE ) MIDWEST RES INST
                                                                                                                                                                        Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO2003012090-A2
                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Acidothermus
                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-FEB-2003
                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                   18
                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                                                                           331
                                                                                                                                                                                                                                                                                                                                                                                                                                      ABZ77632
                                                                                                                                                                                  Query Match:
                                                                                                                                            ..
92
                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 4
 88866666666688888
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                   à
```

```
2686 Gradaracegegregregregregradartrefregaegereaecegreegracegreeger 2745
                                                                                                                                                                                                                                                                                 The present sequence encodes a thermostable avicelase polypeptide, designated AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed production of detergents, pulp and parer processing, food and feed useful in the degradation of cellulose, and in generating specific antipolypeptides from genetically enginered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polynucleotide is useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase; ds; gene.
                                                                                                              for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                      useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28 ValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrpPheThr
                                                                                      New thermostable AviIII peptide from Acidothermus cellulolyticus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2869
36
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cellulolyticus avicelase AviIII DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x ABZ77632 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
1. .2868
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product= "AviIII"
                                                                                                                                                                                                                                     Claim 3; Page 24; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADD22922 standard; DNA; 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-OCT-2002; 2002US-00155400.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUL-2001; 2001US-00917376.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9.28e-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36.00
100.00%
100.00%
40.45%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
WPI; 2003-248177/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                P-PSDB; ABP73015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Acidothermus
```

```
The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents DNA encoding the Acidothermus cellulolyticus
                                                                                                                                                                                                New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                    Ding S, Adney WS, Vinzant TB, Himmel ME;
                                                                                                                                                                                                                                                                              Claim 17; SEQ ID NO 2; 29pp; English.
DING S.
ADNEY W S.
VINZANT T B.
(DING/) DING S.
(ADNE/) ADNEY W S.
(VINZ/) VINZANT T B.
(HIMM/) HIMMEL M E.
                                                                                                                                       WPI; 2003-810853/76.
P-PSDB; ADD22921.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    avicelase AviIII
```

Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other; Length:
Matches:
Conservative:
Mismatches:
Indels: 9.28e-26 36.00 100.00% 100.00% 40.45% Percent Similarity: Best Local Similarity: Alignment Scores: Pred. No.: Query Match: DB:

US-09-917-376-4 (1-89) x ADD22922 (1-2869)

ð g

ઠે

Search completed: May 12, 2004, 15:18:40 Job time : 324 secs

```
May 12, 2004, 15:15:37 ; Search time 1965 Seconds (without alignments) 1352.536 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                               GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                   OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27513289 segs, 14931090276 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Listing first 1000 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Xgapop 60.0 , Xgapext 60.0 Ygapop 60.0 , Ygapext 60.0 Fgapop 6.0 , Fgapext 7.0 Delop 6.0 , Delext 7.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fun: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inv:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_gss_pln:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mam: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_gas_mus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gss_rod:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_gss_phg:*
em_gss_vrl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_estba:*
em_esthum:*
em_esthum:*
em_esthun:*
em_estru:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro:*
em_estro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em estfun:*
                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_estom:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gb_est5:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em gss
em gss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em gss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OFIGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EST:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30
                                                                                                                                                                                                                                                                                                                                                                               Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Word size:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                               Run on:
```

29: gb gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description Query Score Match Length DB Result No.

No matches found

Search completed: May 12, 2004, 16:38:30 Job time : 1968 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
May 12, 2004, 15:18:47 ; Search time 67 Seconds (without alignments) 737.174 Million cell updates/sec
                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                         682709 segs, 277475446 residues
                                                                                                                                                                                    5 60.0 , Xgapext 60.0
5 60.0 , Ygapext 60.0
5 6.0 , Fgapext 7.0
6.0 , Delext 7.0
                                                                                               US-09-917-376-4
89
                                                                                                                                                                                                       Ygapop
Fgapop
Delop
                                                                                                                                                                                      Xgapop
                                                                                                                                                                      OLIGO
                                                                                               Title:
Perfect score:
Sequence:
                                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                                                           Searched:
                                 Run on:
```

Post-processing: Listing first 1000 summaries

0

Total number of hits satisfying chosen parameters:

30

Word size:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

**MODEL=frame+ parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter parameter p Command line parameters:

/cgn2_6/ptodata/2/ina/5A_COMB.seq:*
/cgn2_6/ptodata/2/ina/5B_COMB.seq:*
/cgn2_6/ptodata/2/ina/6A_COMB.seq:*
/cgn2_6/ptodata/2/ina/PcTUS_COMB.seq:*
/cgn2_6/ptodata/2/ina/PcTUS_COMB.seq:*
/cgn2_6/ptodata/2/ina/PcTUS_COMB.seq:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Score Match Length DB Result Š.

Description

No matches found

Search completed: May 12, 2004, 16:39:50 Job time : 69 secs

Appli Appli Appli Appli Appli

Title: Perfect score:

Sequence:

protein

Run on:

Scoring table:

Word size:

Searched:

```
1543 TGGGCGGCGATGGGGGTGTGGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
Sequence 2, 1
Sequence 2, 1
Sequence 2, 1
Sequence 2, 1
Sequence 2, 1
                                                                                                                                                                                                                                                                                                                                                                               Sequence 2, Application US/09917378
; Sequence 2, Application US/09917378
; Sequence 2, Application US/09917378
; Sequence 2, Application US. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: VINZANT, TOOD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMEL TOLERANT MANNANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE OF INVENTION: CELLULOLYTICUS
; TILLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2289
88
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
Indels:
Gaps:
US-09-917-378-2
US-09-917-384-2
US-09-917-383-2
US-09-917-376-2
US-10-155-400-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-4 (1-89) x US-09-917-378-2 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1603 ccgacgedededecrecerged 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TYPE: DNA; ORGANISM: Acidothermus cellulolyticus US-09-917-378-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
US-09-917-384-2
IS-09-917-384-2
; Sequence 2, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: VINZANT, TODD B.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
           12010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.3e-82
88.00
100.00$
100.00$
98.88$
       2289
3687
3687
2869
2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
   98.9
93.3
40.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
Pred. No.:
           88
83
36
36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OCENTIAL ALTONOME AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF A THROUGH AND ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF ACT OF A
                                                                                                                                                                                                                                   ; Search time 347 Seconds
  (without alignments)
  1161.871 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "(gn2_6/ptodata1/1/pubpna/US07_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US06_MMF_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US06_MMF_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US06_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US07_NEW_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US07_NEW_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US08_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US08_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US08_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US08_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US08_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US09_NEW_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US09_NEW_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US09_NEW_PUB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
"(gn2_6/ptodata1/1/pubpna/US10A_PUBCOMB.seq:*)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX
       version 5.1.6
- 2004 Compugen Ltd
                                                                                                                                                       using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2941586 segs, 2264995651 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Listing first 1000 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Xgapext 60.0
Ygapext 60.0
Fgapext 7.0
Delext 7.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Published_Applications_NA:*
                                                                                                                                                                                                                                       May 12, 2004, 16:39:58
       GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-4
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60.0
                                                                                                                                                               nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Xgapop 6
Ygapop 6
Fgapop
Delop
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OFIGO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30
```

Jatabase :

9

Score

Result

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
US-09-917-376-2
                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTCACCCGGGATGGTGGGTCGTCGTCGTCGTGTACAACTGTGACTGGGCGGCGATGGGG 1944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1945 TGTGGGAATATCCGCGCCTCGTTCGGCTGGTGAACCCGGCGACGCCGGGGGACACC 2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1765 AAGGIGCAGIACAAGAACAAIGAIICGCCCCCCGGIGAIAACCAGAICAAACCGGGICIC 1824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1825 CAGTTGGTGAATACCGGGTCGTCGTCGTGGATTTGTCGACGGTGACGGTGCGGTACTGC 1884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyrTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ## Sequence 2, Application US/09917383

## Sequence 2, Application US/09917383

## Publication No. US20030104522A1

## SEQUENCE 2, SEQUENCE 3, SEPTION S.

## APPLICANT: DIGKER, STEPHEN R.

## APPLICANT: DECKER, STEPHEN R.

## APPLICANT: HIMMEL, MICHAEL B.

## APPLICANT: HIMMEL, MICHAEL B.

## TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS

## TITLE OF INVENTION: CELLULOLYTICUS

## TITLE OF INVENTION: CELLULOLYTICUS

## TITLE OF INVENTION: CELLULOLYTICUS

## TITLE OF INVENTION: CELLULOLYTICUS

## TITLE OF INVENTION: CELLULOLYTICUS

## TITLE OF INVENTION UNMBER: US/09/917,383

## CURRENT FILING DATE: 2001-07-28

## SOFTWARE: PatentIn Ver. 2.1

## SEQ ID NOS: 14

## SEQ ID NOS: 14
                                                                                                                                                                                                                                                                OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION; GuxA
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170.6UGUS
CURRENT APPLICATION UNDBER: US/09/917,384
CURRENT APPLICATION UNDBER: US/09/917,384
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PALENTIN Ver. 2.1
LENGTH: 3687
                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-376-4 (1-89) x US-09-917-384-2 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                          1.09e-76
83.00
100.00%
100.00%
93.26%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |||||||||
2005 TACCTGCAG 2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 3687
                                                                                                                                                                                                                                                                                                             US-09-917-384-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-383-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

Alignment Scores:

```
2686 GTGAATACCGGGTCGTCGTCGGTGTTTGTCGACGGTGACGGTGCGGTACTGGTTCACC 2745
                                                                                                                                                                                                                    1825 CAĞTTGGTGAATACCGGGTCGTCGTCGGTGGATTTGTCGACGGTGACGGTGCGGTACTGG 1884
                                                                                                                                                                                                                                                                                                                                                                                              45
                                                                                                                                                                                                                                                                                                                                                                    46 PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28 ValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrpPheThr 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/09917376

Sequence 2, Application US/09917376

Publication No. US2004003834A1

GENERAL INFORMATION:

APPLICANT: DING, SHILLIAM S.

APPLICANT: VINZANT: TOOD B.

APPLICANT: HIMMEL, MICHAEL E.

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: THERMAL TOLERANT

SOFTWARE: PALENTIN VET. 2.1

SOFTWARE: PALENTIN VET. 2.1

SEQ ID NO 2

LENGTH: 2869
                                                                                                                                                                                              6 LysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu
                                                                                                                                                                                                                                                                                  26 GlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2746 CGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGACTGGGCGGCG 2793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             48 ArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2869
36
0
0
3687
83
0
0
0
0
Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           other or unknown
                                                                                                                                                    US-09-917-376-4 (1-89) x US-09-917-383-2 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-376-4 (1-89) x US-09-917-376-2 (1-2869)
                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
1.09e-76
83.00
100.00%
100.00%
93.26%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.08e-27
36.00
100.00$
100.00$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEX: modified_base
LOCATION: (2869)
CTHER INFORMATION: a, c, t,
US-09-917-376-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2005 raccrecae 2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86 TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                            Percent Similarity:
Best Local Similarity:
```

```
RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 5

RESULT 6

RESULT 6

RESULT 7

RESULT 6

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT 7

RESULT
```

Search completed: May 12, 2004, 17:28:07 Job time : 349 secs

```
ADD22925
                                                                                                                                                                                                    Amino aci
Amino aci
Amino aci
Amino aci
Amino aci
                                                                                                                                                                                                                                                                                                                                                                                                                             A. cellul
A. cellul
Cellulase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protein e
Protein e
Bacillus
                                                                                                                                                                                                                                                                                                                                                    Acidother
Acidother
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Active ce
Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Active ce
Truncated
                                                                                                                                                                                                                                                                                                                                                                    Amino aci
Amino aci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NK-1 cell
Novel end
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Active ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Corrected
                                        May 11, 2004, 11:39:45; Search time 14.4613 Seconds (without alignments) 1738.900 Million cell updates/sec
                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                              470
1 VSGGVKVQYKNNDSAPGDNQ......RASFGSVNPATPTADTYLQX 89
                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                    Add22925 Add22924 Add22924 Abp73022 Abp73022 Abp73018 Abp73018 Abp73015 Abp73015 Abp71658 Aar15652 Aar13494 Aae16323 Aay13492 Aae16323 Aay13492 Aae16323 Aay13492 Aae16323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Aao22444
Aao22443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Aau98063 |
Aar42122 |
Aar13227 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Aaw18790
                                                                                                                                          1586107
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                           1586107 seqs, 282547505 residues
                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                             - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                 ABP73025
ABP73029
ABP73018
ABP73017
ABP73015
                                                                                                                                                                                                                                                                                                                                                                                                                             ABP71658
ABP71656
AAR15625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAR42122
AAR13227
AAW18790
                                                                                                                                                                                                                                                                                                                                                    ADD22925
ADD22924
                                                                                                                                                                                                                                                                                                                                                                  ABP73020
ABP73022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY13493
AAE16324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAE16323
AAO22444
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY13494
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAE16325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY13492
                                                                                                            Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                      A Geneseq 29Jan04:*
                                                                                                                                                                                                                                                geneseqp2002s:*
geneseqp2003as:*
                                                                                                                                                                                                                                                              geneseqp2003bs:*
                                                                                                                                                                                                                        geneseqp1990s:*
                                                                                                                                                                                                                                        geneseqp2001s:*
                                                                                                                                                                                                                  geneseqp1980s:*
                                                                                                                                                                                                                                 geneseqp2000s:*
                                                                                                                                                                                                                                                                     geneseqp2004s:*
                                                                                                                                                        seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                       В
                                                                        US-09-917-376-4
                                                                                                                                                                                                                                                                                                                              Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                  616
616
1751
1751
1426
1426
477
                                                                                                     BLOSUM62
                                                                                                                                                                                                                                                                                                                                                    40.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40.7
                                                                                                                                                                                                                                                462
436
231.5
195.5
195.5
194.5
191.5
188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   186
172
165.5
                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                           Perfect score:
                             OM protein
                                                                                                                                                        Minimum DB
Maximum DB
                                                                                       Sequence:
                                                                                                                           Searched:
                                                                                                                                                                                                           Database
                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                            Result
No.
```

Aae05746 Clostridi	AAE05746	4	190	24.0	113	45
_	AAE05748	4	428	24.3	114	44
-	AAE05747	4	341	24.3	114	43
_	AAE05749	4	328	24.3	114	42
_	AAU97869	S	256	24.3	114	41
~	AAB81128	4	256	24.3	114	40
~	AAE05745	4	163	24.3	114	33
Aaw90077 C. cellul	AAW90077	7	162	24.3	114	38
Aar63634 Cellulose	AAR63634	7	162	24.3	114	37
Aar13229 Endogluca	AAR13229	7	532	25.6	120.5	36
Aaw01503 60 kD end	AAW01503	7	531	25.6	120.5	35
Abu23559 Protein e	ABU23559	9	1483	27.8	130.5	34
-	AAG63963	4	1350	32.0	150.5	33
Aag63962 Amino aci	AAG63962	4	1352	33.1	155.5	32
Aaw43108 C. thermo	AAW43108	~	1853	35.1	165	31
Aaw15238 Scaffoldi	AAW15238	~	531	35.1	165	30
Aay43218 Pectate l	AAY43218	7	493	35.1	165	53
	AAY28850	7	493	35.1	165	28
Aay54123 A mannana	AAY54123	m	476	35.1	165	27
Aar95080 Cellulose	AAR95080	~	167	35.1	165	56

ALIGNMENTS

enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase. Acidothermus cellulolyticus avicelase AviIII CBD III #2. Himmel ME; ADD22925 standard; protein; 88 AA. Adney WS, Vinzant TB, 28-JUL-2001; 2001US-00917376 18-OCT-2002; 2002US-00155400 Acidothermus cellulolyticus. (first entry) m. WPI; 2003-810853/76. (VINZ/) VINZANT T B (HIMM/) HIMMEL M E. DING S. ADNEY W S. VINZANT T B US2003108988-A1. 15-JAN-2004 12-JUN-2003. ADD22925; (DING/) (ADNE/) Ding S,

New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.

Claim 16; SEQ ID NO 5; 29pp; English.

The The invention relates to an isolated polynucleotide molecule encoding a thermostable Avilli polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding Avilli. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase Avilli CBD III #2. Indels

```
Matches
                                                                                                                                  ABP73020
                                                                                                                                                                g
                                          셤
                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII CBD III #1.
                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                       VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                          1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                     Gaps
                                                    ;
0
                                                                                                                                                                                                                                                                                    enzyme, AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
                               DB 7; Length 88;
                                                    Indels
                                                                                                                                                                                                                                                                Acidothermus cellulolyticus avicelase AviIII CBD III #1.
                              Query Match 99.8%; Score 469; DB 7; L
Best Local Similarity 100.0%; Pred. No. 3.9e-46;
Matches 88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Himmel ME;
                                                                                                                     WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                          WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 16; SEQ ID NO 4; 29pp; English
                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                 /label= Unknown
                                                                                                                                                                                               ADD22924 standard; protein; 89
                                                                                                                                                                                                                                                                                                                                                                                                                                18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                     Acidothermus cellulolyticus
                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (DING/) DING S.
(ADNE/) ADNEY W S.
(VINZ/) VINZANT T B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2003-810853/76.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Adney WS,
                                                                                                                                                                                                                                                                                                                                            Key
Misc-difference
           Š
                                                                                                                                                                                                                                                                                                                                                                                       US2003108988-A1
           Sequence 88
                                                                                                                                                                                                                                           15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                          12-JUN-2003
                                                                                                                                                                                                                     ADD22924;
                                                                                                                      61
                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ding
                                                                                                                                                                                     ADD2292
                                                                                                                                                                           RESULT
                                                                                            g
                                                                                                                                                                                                ò
                                                                                                                     ઠે
                                                                                                                                      g
```

Length 89;

99.8%; Score 469; DB 7; 100.0%; Pred. No. 4e-46;

Best Local Similarity

Query Match

Sequence 89 AA;

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present sequence represents a fragment of ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                9
                                                                                              9
                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar; feed; paper pulp; biofuel; mannase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Amino acid sequence of the ManA carbohydrate binding domain type III
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Indels
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99.8%; Score 469; DB 6; I 100.0%; Pred. No. 7.8e-46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%; Pred. ....
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                 88
                                                                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Himmel
                                                                                                                                                 61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                          Ą
                                                                                                                                                                                                                                                                                                                        ABP73020 standard; peptide; 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŢB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 8; Page 7; 46pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2003-248182/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 88; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ManA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 154 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO2003012110-A1
88;
                                                                                                                                                                                                                                                                                                                                                                                                                         03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ManA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-FEB-2003
                                                                       13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABP73029;
                                                                                                                                                           s,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABP73029
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ⋧
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                The present sequence represents ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicallulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA
                                                                                                                                                                                                                                                                                                                  Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            455 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                     ManA; mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar; food; feed; paper pulp; biofuel; mannase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Amino acid sequence of the GuxA carbohydrate binding domain type III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.8%; Score 469; DB 6; Length 762;
100.0%; Pred. No. 5.7e-45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Indels
                                                                                 sequence of the ManA polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Preq. ...
                                                                                                                                                                                                                                                                  Himmel ME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                    ABP73022 standard; protein; 762 AA.
                                                                                                                                                                                                                                                                                                                                                                Claim 3; Page 18-19; 46pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABP73025 standard; peptide; 150
                                                                                                                                                                                                                                                                 Vinzant TB,
                                                                                                                                                                                                  28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                        28-JUL-2001; 2001WO-US023819.
                                                                                                                                     Acidothermus cellulolyticus
                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                            (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                      WPI; 2003-248182/24
                                                                                                                                                                                                                                                                  Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sest Local Similarity
                                                                                                                                                                                                                                                                                                 N-PSDB; ABZ77633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 762 AA;
                                                                                                                                                          WO2003012110-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        03-JUN-2003
                                                            03-JUN-2003
                                                                                                                                                                                13-FEB-2003.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88;
                                                                                  Amino acid
                                       ABP73022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP73025
                                                                                                                                                                                                                                                                 Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
ABP73025
ID ABP7
XX
AC ABP7
XX
DT 03-J
XX
DE Amin
XX
                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

```
Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABP73023-28 represent peptide fragments of a GuxA polypeptide. GuxA is a thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biotuel production. GuxA is also useful in the conversion of biomass to biotuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
cellulose; sugar; biofuel; feed processing; textile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cellulose; sugar; biofuel; feed processing; textile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGLKVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 466; DB 6;
Pred. No. 1.7e-45;
GuxA; cellulase; glycoside hydrolase; enzyme;
detergent; pulp processing; paper processing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GuxA; cellulase; glycoside hydrolase; enzyme;
detergent; pulp processing; paper processing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amino acid sequence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP73029 standard; protein; 1228 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 6; Page 7; 47pp; English.
                                                                                                                                                                                                                                                                   28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99.1%;
98.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Acidothermus cellulolyticus
                                                                                      Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 98.9
Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                        (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2003-239526/23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ws,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 150 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO2003012109-A1
                                                                                                                                                WO2003012109-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Adney
```

```
g
                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                      Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                        The present sequence represents a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                   584 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 643
                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                       VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Amino acid sequence of AviIII carbohydrate binding domain type II.
                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                         99.1%; Score 466; DB 6; Length 1228; 98.9%; Pred. No. 2.3e-44;
                                                                                                                                                                                                                                                                                                                               0; Indels
                                  Decker SR;
                               Vinzant TB, Himmel ME,
                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                 644 WAAMGCGNIRASFGSVNPATPTADTYLQ 671
                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABP73018 standard; peptide; 88 AA
                                                                                                                                    Claim 3; Page 19; 47pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                               87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (MIDE ) MIDWEST RES INST.
           (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2003-248177/24
                                                      WPI; 2003-239526/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Adney WS,
                               Adney WS,
                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                     Sequence 1228 AA;
                                                                  N-PSDB; ABZ77634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                       Н
                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ς,
                                  Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ding
                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                           유
                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
The present sequence is derived from a thermostable avicelase, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a callulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in the processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Avicelase, AviIII, glycoside hydrolase, enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Amino acid sequence of AviIII carbohydrate binding domain type III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 98.3%; Score 462; DB 6; Length 88; Best Local Similarity 97.7%; Pred. No. 2.5e-45; Matches 86; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note= "unspecified residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 WAAIGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABP73017 standard; peptide; 89
                                                                Claim 6; Page 8; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 6; Page 8; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-2001; 2001WO-US023818.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2003-248177/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 88 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   host cells.
host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABP73017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 8
```

```
ö
         designated
The present sequence is derived from a thermostable avicelase, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide tool for characterizing the molecular source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence represents a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Avicelase, AviIII; glycoside hydrolase, enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGSVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRXWFTRDGGSSTLVXNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "unspecified residue encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 462; DB 6; L4
Pred. No. 2.5e-45;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Amino acid sequence of the avicelase AviIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABP73015 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 2; Page 20; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Adney WS, Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ch 98.3%;
1 Similarity 97.7%;
86; Conservative ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2003-248177/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; ABZ77632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 89 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABP73015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP73015
ABP
         8888888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
```

```
ö
biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AvilII peptide is useful in the degradation of cellulose, and in generating specific antipoptides from genetically engineered host cells, in detecting AvilII polypeptides from genetically engineered host cells, in detecting AvilII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AvilII polynucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                            869 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to an isolated polynucleotide molecule encoding thermostable AviII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves
                                                                                                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                           Length 957;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
                                                                                                                                                                                                       Query Match

98.3%; Score 462; DB 6; Length 95
Best Local Similarity 97.7%; Pred. No. 4.8e-44;
Matches 86; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Acidothermus cellulolyticus avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Vinzant TB, Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                926
                                                                                                                                                                                                                                                                                                                                               88
                                                                                                                                                                                                                                                                                                                                                                  929 WAAIGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 16; SEQ ID NO 1; 29pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADD22921 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /label= Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2003-810853/76.
N-PSDB; ADD22922.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADNEY W S.
                                                                                                                                                                             Sequence 957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (DING/) DING S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADD22921;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADNE/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (/ZNIA
                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                              셤
```

ö

Gaps

; 0

Length 153; Indels

Score 436; DB 6; Pred. No. 4.9e-42; 1; Mismatches 2;

Query Match 92.8%; Best Local Similarity 96.4%; Matches 81; Conservative

64 63

5 VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM 4 LKAQYKNNDSAPSDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM

88 87

65 GCGNIRASFGSVNPATPTADTYLO 64 GCGNIRASFGSVNPATPTADTYLO

셤

ABP71656 standard; protein; 1121 AA

RESULT 12

g

δ

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The invention relates to a thermal tolerant Gux1 peptide from A. cellulolyticus. The Gux1 exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type III, and a carbohydrate binding domain type III, and a carbohydrate binding domain type III. The polypeptide is useful in the degradation of cellulose into biofuel, or pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial-scale processing of cellulase to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Gux1 cellulase CBD_III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 869 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYRFTRDGGSSTLVYVCD 928
                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                            The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New thermal tolerant Gux1 peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ermal tolerant; exoglucanase; glycoside hydrolase; cellulose; detergent; pulp; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                Score 462; DB 7; Length 957;
Pred. No. 4.8e-44;
2; Mismatches 0; Indels

    A. cellulolyticus Gux1 protein CBD_III domain fragment.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABP71658 standard; protein; 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 2; Page 7; 44pp; English
                                                                                                                                                                                                                                                                                                       98.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 thermal tolerant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-300494/29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cellulase; enzyme
                                                                                                                                                                                                                                      Sequence 957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO2003012095-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                    86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABP71658;
                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      biofuel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 111
ABP71658
ID ABP71658
ID ABP71658
XX ABP71658
XX GUXLI,
XW Diofulty
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID

      8888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
```

Gux1; thermal tolerant; exoglucanase; glycoside hydrolase; cellulose; biofuel; detergent; pulp; paper processing; feed processing; textile;

A. cellulolyticus Gux1 protein.

(first entry)

29-MAY-2003

ABP71656;

.34
 hote= "potential signal peptide" 35. .1121
 hote= "mature protein"

Location/Qualifiers

Peptide Protein

Acidothermus cellulolyticus.

cellulase; enzyme.

/note= "encoded by CG"

Misc-difference

```
The invention relates to a thermal tolerant Gux1 peptide from A. cellulolyticus. The Gux1 excglucanase is a member of the glycoside hydrolase family and comprises a caralytic domain GH48, carbohydrate binding domain type III. and a carbohydrate binding domain type III. The polypeptide is useful in the degradation of cellulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial-scale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Gux1 cellulase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                                                                                                                                                                                                                                           Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 3; Page 18-19; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                           TB,
                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                                                                                                                                                                                                                                                           Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                             (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-300494/29.
N-PSDB; ABZ76162.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1121 AA
                                                                                                                                                                                                                                                                                           WO2003012095-A1
                                                                                                                                                                                                                                                                                                                         13~FEB-2003
```

Sequence 153 AA;

domain

(first entry)

```
Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel B1; Cel Cel 1/2/3; Cel 6; Cel B3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
           AAY13494 standard; protein; 616 AA.
                                                                                                                                                                                                                                                               (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                             Truncated cellulase Cel E3/B5.
                                                                                                                                                                                                                                                                                    Anderson P, Bergquist PL,
Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                      WPI; 1999-315403/27.
                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAX55660
                                                                                                                                                                                                                  L5-SEP-1998;
                                                                                                                                                                                                                                         19-SEP-1997;
                                                                                                                                                 Unidentified
                                                       30-JUL-1999
                                                                                                                                                                      EP921188-A2
                                                                                                                                                                                            09-JUN-1999
                                 AAY13494;
AAY13494
                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The sequence was deduced from the gene which was sequenced from plasmid, pAEC 1, prepd. by ligating chromosomal DNA contg. the gene (obtd. from Aeromonas) into put18. The protein has amol. wt. of 81,000 (SDS-PAGE) and an optimum pH near to 5 when carboxy—methylcellulose is the substrate. The N-terminal sequence: GIHADT—has been confirmed by Edman degradation. The gene can be used to produce recombinant enzyme which is used for the effective utilis—ation of biomass resources and the mfr. of pharmacceuticals and foodstuffs, and also for the deterrgent and deinking of waste paper. (Updated on 25-MAR-2003 to correct PA field.)
                                                        64
                                                                             97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   opt.
                                                                   38 LKAQYKNNDSAPSDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM
                                                     5 VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM
                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cellulase AB-1 for e.g. mfr. of pharmaceuticals and foodstuffs - of pH when carboxy:methyl:cellulose is used as substrate.
                                 ö
                                                                                                                                                                                                                                                                                     Detergents; pharmaceuticals; deinking; carboxymethylcellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49.3%; Score 231.5; DB 2; Length 782; 48.3%; Pred. No. 1.3e-17; ive 16; Mismatches 28; Indels 1
          Length 1121;
            Score 436; DB 6; Length 11
Pred. No. 5.7e-41;
1; Mismatches 2; Indels
                                                                                                                 GCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                  AAR15625 standard; protein; 782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 2; Fig 3; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                              90JP-00045465
                                                                                                                                                                                                                                                                                                                                                                                                   90JP-00045465
           92.8%;
96.4%;
                                                                                                                                                                                                                              (revised)
(first entry)
                                   81; Conservative
                                                                                                                                                                                                                                                                                                           Aeromonas strain no. 212
                                                                                                                                                                                                                                                                                                                                                                                                                        (OJIP ) OJI PAPER CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1991-373412/51.
                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-PSDB; AAQ15178.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 782 AA;
                                                                                                                                                                                                                                                               Cellulase AE-1
                                                                                                                                                                                                                                                                                                                                                                             28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                   28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                  JP03251174-A.
                                                                                                                                                                                                                             25-MAR-2003
17-MAR-1992
                                                                                                                                                                                                                                                                                                                                                        08-NOV-1991,
                                                                                                                                                                                                       AAR15625;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
           Query Match
Best Local
                                                                                                   65
                      Best Loca
                                                                                                                                                             RESULT 13
                                                                                                                                                                      g
                                                                                                   ò
                                                                                                                         셤
```

Gibbs MD;

Farrington GK,

Daniels RM,

97US-00932571. 98EP-00810919

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1/2, Cel 1/2/3, Cel 6 or Cel B3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A101 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid K335 to N1426 in the sequence shown in AAY1342; Cel E1 extends from amino acid K335 to N1426 in the sequence shown in AAY1342; Cel E1/2/3 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability backstaining detergent compositions acid 6482 to G635 in the stabiling cotton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewaching, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using nontruncated callulase compositions
New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41.6%; Score 195.5; DB 2;
43.2%; Pred. No. 1.3e-13;
tive 15; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 WAQIGASNVTFNFVKLSSGVSGADYYLE
                                                                                                                             Claim 7; Page 42-43; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 43.28
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 616 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

AAE16325 standard; protein; 616 AA. RESULT 15 AAE16325 ID AAE1

RESULT 14

AMVĠĊSNINGAFVKMNPGKANAĎIÝĽE 716 62 AAMGCGNIRASFGSVNPATPTADTYLQ 88

069

2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW

Local Similarity 48.3 nes 42; Conservative

Best Loc Matches

a à

8

61

ij

AAE16325;

```
The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the CelBS amino acid sequence. The cellulase active protein consists of treating cellulosic materials including cotton-containing fabrics, as improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing kits or for imparting stonewashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present sequence is E3/B5 hybrid protein
                                                                                                                                                                                                                                                                Active cellulase protein; alkalinophilic; textile processing; proteinase; detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; E3/B5 hybrid protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bergquist P, Daniels R, Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41.6%; Score 195.5; DB 5; Length 616; 43.2%; Pred. No. 1.3e-13; ative 15; Mismatches 34; Indels 1.
                                                                                                                                                                                               Active cellulase hybrid protein, E3/BS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Col 75-78; 61pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-00136574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97US-00932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Farrington GK, Anderson P,
Morgan H, Williams DP;
                                                                                                                 26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       elevated temperatures.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2002-081780/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
nes 38; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAD26568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 616 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                              Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US6294366-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac{1}{2}$\frac
```

1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60 Gaps 1; Conservative ઠે g

WAAMGCGNIRASFGSVNPATPTADTYLQ 88

ò

요

Search completed: May 11, 2004, 12:06:51 Job time : 16.4613 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

protein search, using sw model • OM protein May 11, 2004, 11:39:45; Search time 14.2988 Seconds (without alignments) 1738.900 Million cell updates/sec Run on:

US-09-917-376-5

Perfect score:

1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ 88 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

1586107 seqs, 282547505 residues Searched:

1586107 Total number of hits satisfying chosen parameters:

seq length: 0 seq length: 200000000 Minimum DB & Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

A Geneseq 29Jan04:* geneseqp1980s:* Database

geneseqp2003as:* geneseqp2003bs:* geneseqp2001s:* geneseqp2002s:* geneseqp1990s:* geneseqp2000s:*

geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	Add22925 Acidother	Add22924 Acidother	Abp73020 Amino aci	Abp73022 Amino aci	Abp73025 Amino aci	Abp73029 Amino aci	Abp73018 Amino aci	Abp73017 Amino aci	Abp73015 Amino aci	Add22921 Acidother	Abp71658 A. cellul	Abp71656 A. cellul	Aar15625 Cellulase	Aay13494 Truncated	Aae16325 Active ce	Aay13493 Truncated	Aael6324 Active ce	Aay13492 Truncated	Aae16323 Active ce	Aao22444 Protein e	Aao22443 Protein e	Aau98063 Bacillus	Aar42122 NK-1 cell	Aarl3227 Novel end	Aaw18790 Corrected
Sa LANGUIGO	ID	ADD22925	ADD22924	ABP73020	ABP73022	ABP73025	ABP73029	ABP73018	ABP73017	ABP73015	ADD22921	ABP71658	ABP71656	AAR15625	AAY13494	AAE16325	AAY13493	AAE16324	AAY13492	AAE16323	AA022444	AA022443	AAU98063	AAR42122	AAR13227	AAW18790
	DB	7	7	9	9	9	9	9	9	9	7	9	9	∾	N	Ŋ	~	ហ	N	ហ	ഗ	'n	ហ	7	~	N
	Query Match Length	88	89	154	762	150	1228	88	8	957	957	153	1121	782	616	616	1751	1751	1426	1426	477	496	677	499	700	551
dР	Query Match	100.0	100.0	100.0	100.0	P. 66	99.4	98.5	98.5	98.5	98.5	93.0	93.0	49.4	41.7	41.7	41.5	41.5	40.8	40.8	40.1	40.1	40.1	39.7	36.7	35.3
	Score	469	469	469	469	466	466	462	462	462	462	436	436	231.5	195.5	195.5	194.5	194.5	191.5	191.5	188	188	188	186	172	165.5
	Result No.	п	7	m	4)) ′	y	7	œ	σ	10	11	12	Ŀ	14	15	16	17	18	19	20	21	22	23	24	25

Aar95080 Cellulose Aay54123 A mannana Aay58850 Pectate 1 Aay43218 Pectate 1 Aaw43108 C. thermo Aag63962 Amino aci Abu23559 Protein e Aaw01503 60 kD end Aar13229 Endogluca Aar3229 Endogluca Aar63634 Cellulose Aaw90077 C. cellul Aae05745 Clostridi Aab81128 C1762 Osp Aau97869 B. Coli C	
AAR95080 AAY54123 AAY4218 AAY4318 AAW15238 AAW13108 AAG63962 AAG63962 AAG63963 AAR13229 AAR13229 AAR63634 AAR81286	AAE05747 AAE05747 AAE05748 AAE05746
0 m 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	† ব ব ব
1677 1083 1083 1083 1085 1085 1085 1085 1085 1085 1085 1085	341 428 190
23 25 25 25 25 25 25 25 25 25 25 25 25 25	244.3 24.3 24.3 24.3
	111111111111111111111111111111111111111
00000000000000000000000000000000000000	4 4 4 4 2 6 4 0

ALIGNMENTS

enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase. Acidothermus cellulolyticus avicelase AviIII CBD III #2. ADD22925 standard; protein; 88 AA (first entry) 15-JAN-2004 ADD22925; RESULT 1 ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 29 25

ADD2 25

ADD2 25

ADD2 25

ADD2 25

ADD2 25

ADD2 25

ADD2 25

ADD

Acidothermus cellulolyticus.

18-OCT-2002; 2002US-00155400. US2003108988-A1. 12-JUN-2003.

DING S. ADNEY W S. VINZANT T B. (VINZ/) VINZANT T B (HIMM/) HIMMEL M E. (DING/) (ADNE/)

28-JUL-2001; 2001US-00917376.

Himmel ME; Vinzant TB, Ding S, Adney WS,

WPI; 2003-810853/76.

New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.

Claim 16; SEQ ID NO 5; 29pp; English.

The The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucleotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural blomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII CBD III #2.

```
ABP73020;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local 9
                                                                                                                                                                                                                                       food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
Matches
                                                                                                                    RESULT 3
                                                                                셤
                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The invention relates to an isolated polynucleotide molecule encoding a thermostable AviIII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucleotide is useful for reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus cellulolyticus avicelase AviIII CBD III #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                            VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                   Gaps
                                                   .;
0
                             Length 88;
                                                                                                                                                                                                                                                                                 enzyme, AviIII; cellulose reduction, agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
                                                  Indels
                                                                                                                                                                                                                                                           Acidothermus cellulolyticus avicelase AviIII CBD III #1.
                         100.0%; Score 402; 22
100.0%; Pred. No. 3.9e-46;
rive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Vinzant TB, Himmel ME;
                                                                                                                    61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                             Claim 16; SEQ ID NO 4; 29pp; English.
                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                              Ą
                                                                                                                                                                                            ADD22924 standard; protein; 89
                                                                                                                                                                                                                                                                                                                                                                                                                           18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                  Acidothermus cellulolyticus
                                                                                                                                                                                                                                      (first entry)
                                                  88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2003-810853/76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (DING/) DING S. (ADNE/) ADNEY W S. (VINZ/) VINZANT T I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Adney WS,
                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                Misc-difference 89
         Sequence 88 AA;
                                                                                                                                                                                                                                                                                                                                                                                US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 89 AA;
                                                                                                                                                                                                                                       15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                      12-JUN-2003.
                                                                                                                                                                                                                  ADD22924;
                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ding S,
                                                                                                                                                                                             ď
                                                                                                                  ò
                                                                                                                                      요
```

Length 89;

Score 469; DB 7 Pred. No. 4e-46;

100.0%;

Query Match Best Local Similarity

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present sequence represents a fragment of ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                       9
                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 VSGGVKVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
feed; paper pulp; biofuel; mannase.
  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Amino acid sequence of the ManA carbohydrate binding domain type III
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 154;
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 469; DB 6;
Pred. No. 7.8e-46;
; Mismatches 0;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88
                                                                                                                                                                88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vinzant TB,
ö
                                                                                                                                                                                                                                                                                                                                                             ABP73020 standard; peptide; 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 8; Page 7; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus.
88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2003-248182/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ManA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 154 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO2003012110-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ద
```

```
cellulase; glycoside hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 150 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO2003012109-A1
                                                                                                          WO2003012109-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABP73029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence represents ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicallulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and boffenels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VSGGVKVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                 ManA; mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Amino acid sequence of the GuxA carbohydrate binding domain type III
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 762;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 469; DB 6;
100.0%; Pred. No. 5.7e-45;
ive 0; Mismatches 0;
                                                                                                                                                                         sequence of the ManA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             feed; paper pulp; biofuel; mannase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                              ABP73022 standard; protein; 762 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 3; Page 18-19; 46pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABP73025 standard; peptide; 150 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023819.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2003-248182/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; ABZ77633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 762 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            03-JUN-2003
                                                                                                                              03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               515
                                                                                      ABP73022;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ding S,
                                                                                                                                                                                                                                            food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
RESULT 4
ABP73022
ID ABP73022
ID ABP73022
XX ABP73022
XX ABP73022
XX ACID
XX MADA
XX MADA
XX MADA
XX MO2
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX ACID
XX AC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABP73025
ID ABP7
XX
AC ABP7
XX
DT 03-3
XX
XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
```

```
ABP73023-28 represent peptide fragments of a GuxA polypeptide. GuxA is a thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, catecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GuxA; cellulase; glycoside hydrolase; enzyme; cellulose; sugar; biofuel; detergent; pulp processing; paper processing; feed processing; textile.
cellulose; sugar; biofuel; feed processing; textile.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 466; DB 6; Length 15
Pred. No. 1.7e-45;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Decker SR;
      GuxA; cellulase; glycoside hydrolase; enzyme;
detergent; pulp processing; paper processing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino acid sequence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABP73029 standard; protein; 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 6; Page 7; 47pp; English.
                                                                                                                                                                                                                                                                                                             28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023817.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Acidothermus cellulolyticus.
                                                                                                       Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-239526/23.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding S, Adney WS,
```

host cells.

```
ö
                                                                                                                              Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, and for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               584 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNCD 643
                                                                                                                                                                                                                         The present sequence represents a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel;
detergent; pulp processing; paper processing; feed processing; textile;
cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Amino acid sequence of AviIII carbohydrate binding domain type II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 466; DB 6; Length 1228;
Pred. No. 2.3e-44;
1; Mismatches 0; Indels
                                                Decker SR;
                                                Vinzant TB, Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                              Claim 3; Page 19; 47pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABP73018 standard; peptide; 88
                                                                                                                                                                                                                                                                                                                                                                                                                                         99.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87; Conservative
                (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (MIDE ) MIDWEST RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2003-248177/24
                                                                               2003-239526/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Adney WS,
                                                Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1228 AA;
                                                                                            N-PSDB; ABZ77634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABP73018;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ding S,
                                               Ding S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   임
```

```
ö
                                                            designated
                                                                    Aviil: Aviili is a member of the glycoside hydrolase family of enzymes, and is a cellulase. Aviili is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and pare processing, food and feed processing and in textile processes. The thermostable Aviili peptide is useful in the degradation of cellulose, and in generating specific anti-Aviili antibodies that are useful in purifying recombinant Aviili polypeptides from genetically engineered host cells, in detecting Aviili polypeptide actions of the polypeptide. The Aviili polymertide actions of the polypeptide in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amino acid sequence of AviIII carbohydrate binding domain type III.
                                                                                                                                                                                                                                                                                                                        .
0
                                                            The present sequence is derived from a thermostable avicelase,
                                                                                                                                                                                                                                                                                          Length 88;
                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                        Score 462; DB 6; L
Pred. No. 2.5e-45;
2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note≂ "unspecified residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                88
                                                                                                                                                                                                                                                                                                                                                                                                                                             88
                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAIGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TB,
                            Claim 6; Page 8; 44pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABP73017 standard; peptide; 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 6; Page 8; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023818.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 97.7%;
Matches 86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2003-248177/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MIDE ) MIDWEST RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO2003012090-A2
                                                                                                                                                                                                                                                           Sequence 88 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABP73017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nost
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABP73017
임
                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
```

```
ö
       designated
The present sequence is derived from a thermostable avicelase, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in the termostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-AviIII antibodies that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymuclectide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence represents a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is useful in the conversion of
                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                             Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ž
                                                                                                                                                                                                                                                                                               0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "unspecified residue encoded by
                                                                                                                                                                                                                                                            Score 462; DB 6;
Pred. No. 2.5e-45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino acid sequence of the avicelase Avilli
                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Adney WS, Vinzant TB, Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                 88
                                                                                                                                                                                                                                                                                                                                                                                             WAAMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABP73015 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 2; Page 20; 44pp; English.
                                                                                                                                                                                                                                                            98.5%;
97.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03-JUN-2003 (first entry)
                                                                                                                                                                                                                                                                                               86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-248177/24.
N-PSDB; ABZ77632.
                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO2003012090-A2
                                                                                                                                                                                                                            Sequence 89 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABP73015;
                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ding S,
                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
   셤
                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                               δ
```

```
ö
biomass to biofuels and biofuel additives. It may be useful in the production of detergents, pulp and parer processing, food and feed processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific antipolypeptides that are useful in purifying recombinant AviIII polypeptides from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polynucleotide is useful as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                869 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The invention relates to an isolated polynucleotide molecule encoding thermostable AvilII polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding AviIII. The polynucelotide is useful for reducing cellulose in a starting material which involves
                                                                                                                                                                                                                                                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                         Query Match 98.5%; Score 462; DB 6; Length 957; Best Local Similarity 97.7%; Pred. No. 4.8e-44; Matches 86; Conservative 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               enzyme; AviIII; cellulose reduction; agricultural biomass; municipal solid waste; glycoside hydrolase; avicelase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Acidothermus cellulolyticus avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Vinzant TB, Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                       929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                    88
                                                                                                                                                                                                                                                                                                                                                    61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Encoded by N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 16; SEQ ID NO 1; 29pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADD22921 standard; protein; 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /label= Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2003-810853/76.
N-PSDB; ADD22922.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (VINZ/) VINZANT T B
(HIMM/) HIMMEL M E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DING S.
ADNEY W S.
VINZANT T E
                                                                                                                                                                            Sequence 957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US2003108988-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADD22921;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DING/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADNE/)
                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADD22921
   ò
                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                              셤
```

ö

Gaps

..

Length 153;

Score 436; DB 6; Pred. No. 4.9e-42; 1; Mismatches 2

93.0%;

Query Match
Best Local Similarity 96.4
Matches 81; Conservative

64 63

4 LKAQYKNNDSAPSDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM 5 VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM

88

64 GCGNIRASFGSVNPATPTADTYLQ 87

셤 ઠે

65 GCGNIRASFGSVNPATPTADTYLO

원

ABP71656 standard; protein; 1121 AA

RESULT 12 ABP71656

```
869 VSGGVKVQYKNNDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                    9
              The
                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                        Gux1; thermal tolerant; exoglucanase; glycoside hydrolase; cellulose;
biofuel; detergent; pulp; paper processing; feed processing; textile;
administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents the amino acid sequence of Acidothermus
                                                                                                                               Gaps
                                                                                                                              ö
                                                                                                      Length 957;
                                                                                                     Score 462; DB 7; Length 95
Pred. No. 4.8e-44;
2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                 A. cellulolyticus Gux1 protein CBD_III domain fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Himmel ME,
                                                                                                                                                                                                             WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                             ABP71658 standard; protein; 153 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ding S, Vinzant TB,
                                                          cellulolyticus avicelase AviIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-2001; 2001WO-US023820.
                                                                                                     ch 98.5%;
1 Similarity 97.7%;
86; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                            29-MAY-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2003-300494/29.
                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                               cellulase; enzyme.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Adney WS, Ding S,
Lantz Mccarter S;
                                                                               Sequence 957 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                            WO2003012095-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-FEB-2003
                                                                                                                                                                                                                      929
                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                     ABP71658;
                                                                                                     Query Match
                                                                                                                             Matches
                                                                                                                                                                                                                                                       88888888888
                                                                                                                                                    ò
                                                                                                                                                                        셤
                                                                                                                                                                                                ò
                                                                                                                                                                                                                     셤
```

Gux1, thermal tolerant; exoglucanase; glycoside hydrolase; cellulose; biofuel; detergent; pulp; paper processing; feed processing; textile;

A. cellulolyticus Gux1 protein.

(first entry)

29-MAY-2003

ABP71656;

1. .34 /note= "potential signal peptide"

35. .1121

Location/Qualifiers

Peptide Protein

Acidothermus cellulolyticus.

cellulase; enzyme.

```
cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercialscale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Guxi cellulase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The invention relates to a thermal tolerant Guxl peptide from A. cellulolyticus. The Guxl exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type III, and a carbohydrate binding domain type III. The polypeptide is useful in the degradation of cellulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                                                                                                                                                                                                                                              Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Himmel ME,
                                                                                                                                                                                                                        /note= "mature protein" 228
                                                                                                                                                                                                                                                                 /note= "encoded by CG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 3; Page 18-19; 44pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                            Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                     28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                                                                                                                                                                                                                                 (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2003-300494/29.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Adney WS, Ding S,
Lantz Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1121 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; ABZ76162
                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                              WO2003012095-A1
                                                                                                                                                                                                                                                                                                                             13-FEB-2003
```

cellulolyticus. The Guxi exoglucanse is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type III, and a carbohydrate binding domain type III, and a carbohydrate binding domain type III. The polypeptide is useful in the degradation of callulose into biotuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other callulase or glycoside hydrolases. The novel polypeptide generates alternative callulase enzymes capable of assisting in the commercial-scale processing of callulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Guxl cellulase CBD_III

Sequence 153 AA; domain fragment

New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.

invention relates to a thermal tolerant Guxl peptide from A.

Claim 2; Page 7; 44pp; English

```
Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel B1, Cel B1/2; Cel 1/2/3; Cel 6; Cel B3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing.
           AAY13494 standard; protein; 616 AA.
                                                                                Truncated cellulase Cel E3/B5.
                                                                                                                                                                                                                                                   19-SEP-1997;
                                                                                                                                                                                                                            .5-SEP-1998;
                                                         30-JUL-1999
                                                                                                                                                     Inidentified
                                                                                                                                                                                                     09-JUN-1999.
                                                                                                                                                                             EP921188-A2,
                                                                                                                                                                                                                                                                                                               Morgan H,
AAY13494
             Š
                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         pAEC 1, prepd. by ligating chromosomal DNA contg. the gene (obtd. from Ascromona) into pUC18. The protein has amol. wt. of 81,000 (SDS-PAGE) and an optimum pH near to 5 when carboxy—exthylcellulose is the substrate. The N-terminal sequence: GIHADT- has been confirmed by Edman degradation. The gene can be used to produce recombinant enzyme which is used for the effective utilis ation of biomass resources and the mfr. of pharmacceuticals and foodstuffs, and also for the detergent and deinking of waste paper. (Updated on 25-MAR-2003 to correct PA field.)
                                                         64
                                                                              38 LKAQYKNNDSAPSDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The sequence was deduced from the gene whih was sequenced from plasmid,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cellulase AE-1 for e.g. mfr. of pharmaceuticals and foodstuffs - of opt.
pH when carboxy:methyl:cellulose is used as substrate.
                                                       5 VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
                                  ..
                                                                                                                                                                                                                                                                                                 Detergents; pharmaceuticals; deinking; carboxymethylcellulose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         49.4%; Score 231.5; DB 2; Length 782; 48.3%; Pred. No. 1.3e-17; tive 16; Mismatches 28; Indels 1
         Length 1121;
         Score 436; DB 6; Length 11
Pred. No. 5.7e-41;
1; Mismatches 2; Indels
                                                                                                                   GCGNIRASFGSVNPATPTADTYLQ 121
                                                                                                      65 GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                      AAR15625 standard; protein; 782 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 2; Fig 3; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                               90JP-00045465.
         93.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                     90JP-00045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 48.3%
Matches 42; Conservative
                                                                                                                                                                                                                                      (revised)
(first entry)
                                  81; Conservative
                                                                                                                                                                                                                                                                                                                        Aeromonas strain no. 212
                                                                                                                                                                                                                                                                                                                                                                                                                                           (OJIP ) OJI PAPER CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1991-373412/51.
                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAQ15178.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 782 AA;
                                                                                                                                                                                                                                                                        Cellulase AE-1.
                                                                                                                                                                                                                                                                                                                                                                                               28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                     28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                               JP03251174-A.
                                                                                                                                                                                                                                     25-MAR-2003
17-MAR-1992
                                                                                                                                                                                                              AAR15625;
        Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                              Matches
                                                                                                                                                                RESULT 13
                                                                                                                                                                           AAR15625
ID AAR1
                                                                                                                                                                                                   셤
                                                                                                      ઠે
```

98EP-00810919 97US-00932571

(first entry)

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel BS, Cel B4/5, Cel B1/2, Cel 1/2/3, Cel 6 or Cel B3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A39 to D481, Cel E1/2 extends from y39 to G812, Cel E1 E1/2 extends from y39 to G812, Cel E6 extends from y39 to G812, Cel E7 E7 E8 extends from mino acid V1233 to K1751 and the stability region extends from amino acid V1233 to K1751 and the stability region extends from amino acid G812 to G815 in the sequence shown in AAY13493. Cel E3/B5 is shown in AAY13494. The new carymes are useful in laundry detergent compositions to prevent or remove staining, backstaining or graying, for use on cellulosic materials including cotton-containing fabrics. They are especially useful for proventing redeposition of colorant during stonewashing, and for processing of textiles where cellulose breakdown is required. The new truncated enzymes show reduced redeposition of dye compared to using non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                     Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                     Daniels RM, Farrington GK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41.7%; Score 195.5; DB 243.2%; Pred. No. 1.3e-13; ive 15; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 WAQIGASNVTFNFVKLSSGVSGADYYLE 87
                                                                                                                                                                                                                                                                                                                                                                                           Claim 7; Page 42-43; 65pp; English.
(CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        truncated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 45...
38; Conservative ...
                                                                 , Bergquist PL, Williams DP;
                                                                                                                                                                           WPI; 1999-315403/27.
                                                                                                                                                                                                            N-PSDB; AAX55660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 616 AA;
                                                                     Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAE16325
ID AAE1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
```

AAE16325 standard; protein; 616 AA

셤 ð

RESULT 14

```
substantially free of proteiness of critics and alkalinophilic and alkalinophilic origin, where the collulase active protein consists of the CelBs amino acid sequence. The collulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing materials, for stonewashed appearance to cotton-containing whits or for imparting stonewashed appearance to cotton-containing denims. The present proteins surface textile processing and in commercial detergents. The present sequence is B3/Bs hybrid protein
                                                                                                                      Active cellulase protein; alkalinophilic; textile processing; proteinase; detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; E3/B5 hybrid protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present invention relates to a cellulase active protein, which is
                                                                                                                                                                                                                                                                                                                                                                                                                   Bergquist P, Daniels R, Gibbs MD;
                                                                                  Active cellulase hybrid protein, E3/B5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Col 75-78; 61pp; English.
                                                                                                                                                                                                                                                                                                                                                                               (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                           98US-00136574
                                                                                                                                                                                                                                                                                                                                               97US-00932571
                                                                                                                                                                                                                                                                                                                                                                                                                   Farrington GK, Anderson P,
Morgan H, Williams DP;
                                                    26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     elevated temperatures.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2002-081780/11.
N-PSDB; AAD26568.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 616 AA;
                                                                                                                                                                                                                                                                                                           19-AUG-1998;
                                                                                                                                                                                                                                                                                                                                               19-SEP-1997;
                                                                                                                                                                                                  Unidentified
                                                                                                                                                                                                                                      US6294366-B1
                                                                                                                                                                                                                                                                        25-SEP-2001
                AAE16325;
```

Query Match
41.7%; Score 195.5; DB 5; Length 616;
Best Local Similarity 43.2%; Pred. No. 1.3e-13;
Matches 38; Conservative 15; Mismatches 34; Indels 1; Gaps

8 8 8

61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88

60 WAQIGASNVTFNFVKLSSGVSGADYYLE 87

Search completed: May 11, 2004, 12:06:51 Job time : 14.2988 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

May 11, 2004, 11:51:45 ; Search time 4.1265 Seconds (without alignments) 2051.340 Million cell updates/sec

US-09-917-376-5

469 1 VSGGVKVQYKNNDSAPGDNQ......IRASFGSVNPATPTADTYLQ Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283366 Total number of hits satisfying chosen parameters:

283366 seqs, 96191526 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:* 1 2 E 4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	1,4-beta-glucanase	xylanase - Caldice	cellulase (EC 3.2.		cellulase (EC 3.2.	mannan endo-1,4-be	cellulase (EC 3.2.	cellulase (EC 3.2.	cellulase (EC 3.2.	cellulase (EC 3.2.	cellulase homolog	bifunctional cellu	cellulase (EC 3.2.	endoglucanase I (E	cipA protein - Clo	endoglucanase - Er	cellulase (EC 3.2.	thermoactive cellu	scaffolding protei	probably celluloso	scaffolding protei	cellulose 1,4-beta	cellulose-binding	hypothetical prote	_	hypothetical prote	probable RTX famil		
	ID	T31337	T31085	G69593	A26874	JN0111	A48954	T17120	A27198	A43802	S02711	A41897	I40548	B41897	A47704	836859	S39962	S54744	S12021	PC6006	C97012	T30433	S47466	A44140	T08685	T25478	AC2224	B85547	6906	A69196
	DB	7	N	N	N																					~	~	~	7	~
	Length	1711	1779	508	508	499	1331	1742	499	915	1039	145	486	700	879	1854	505	504	986	586	1483	1162	1230	1848	618	547	1428	5188	29	574
* Query	. ~	40.7	ė.	40.5		ö	φ.	39.8	39.7	38.9	38.9	æ	38.4	36.7	35.3	35.2	31.9	Η.	30.5	27.8	7.	7	24.9	4.	16.1			14.9	14.9	14.6
	Score	191	190.5	190	190	188	œ	186.5	186	182.5	182.5	180.5	180	172	165.5	165	149.5	147.5	141.5	130.5	•	'n	117	114	Š.	m.	•	70	7	68.5
Result	No.		7	m	4	S	9	7	ω	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

40.6%; Score 190.5; DB 2; Length 1779;

Query Match

hypothetical prote	thaumatin-like pro	hypothetical prote	ser-lys rich hypot	polyprotein - hepa	probable purH prot	probable reverse g	hypothetical prote	hensin - rabbit	hemolysin - Aquife	sanko - human	endo-beta-1,3-1,4-	replicative DNA he	protein-tyrosine-p	hypothetical prote	Ig heavy chain V r
H72684	JC7201	.T47790	T40462	T01075	C70717	G72614	850669	T30549	H70393	A59386	H95976	AB2512	B44390	AH2515	E37267
7	N	N	~	N	~	N	N	Ŋ	~	~	~	~	Н	~	7
271	247	535	751	1435	523	1222	1571	1594	398	2403	269	454	802	4936	122
14.3	14.2	14.1	14.0	14.0	13.9	13.6	13.6	13.6	13.5	13.5	13.3	13.3	13.3	13.3	13.2
67	66.5	99	65.5	65.5	65	64	64	64	63.5	63.5	62.5	62.5	62.5	62.5	62
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

	RESULT 1
	1,4-beta-glucanase (EC 3.2.1) - Anaerocellum thermophilum (fragment)
	C;Species: Anaeroceilum Chermophilum C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
	Rizerlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K.
	Microbiology 144, 457-465, 1998 Affile: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from
	A. A. E. E. E. C. B. M. M. M. M. M. M. M. M. M. M. M. M. M.
	A;Accession: 13133/ A;Status: preliminary; translated from GB/EMBL/DDBJ
	A;Molecule type: DNA A;Residues: 1-1711 <2VE>
	A;Cross-references: EMBL:Z86105; NID:e1071329; PID:e350354; PIDN:CAB06786.1
	C)Generics: A,Gene: celA
	C;Keywords: glycosidase; hydrolase
	Query Match Best Local Similarity 43.8%; Pred. No. 1.2e-11; Matches 39; Conservative 18; Mismatches 30; Indels 2; Gaps 2;
	QY 1 VSGG-VKVQXKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRXWFTRDGGSSTLVYNC 59
	Db 683 VAGGQIKVLYANKETNSTINTIRPWLKVVNTGSSSIDLSRVTIRYWYTVDGDKAQSAIS- 741
	Qy 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
	,
	Db 742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770
	RESULT 2
	xylangs - Caldicellulosiruptor sp.
	C,Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
	C;Accession: 131085 R;Morris, D.D.; Gibbs, M.D.; Ford, M.; Thomas, J.; Bergquist, P.L.
	submitted to the EMBL Data Library, December 1997 A;Description: Family 10 and 11 xylanase genes from Caldicellulosiruptor sp. Rt69B.1.
	A;Reference number: Z20972
	A;Status: preliminary; translated from GB/EMBL/DDBJ A:Molecule type: DNA
	A; Residues: 1-1779 < MOR>
	A;Cross-references: EMBL:AFU35924; NID:g2760905; FID:g2760909; FIDN:AAB95326.1 C;Genetics:
	A;Note: xynC
_	

G G

Length 508;

```
Residues: 10-13,'V',15,'N',17-18,'V',20-21,'F',23,'A',25-26,'Al',29-31,'P',33,'PO',36-3
TCOSS-references: GB:M38634; NID:G142657; PIDN:AAA22300.1; PID:G142658
Experimental source: strain ATCC 6633
;Comment: The low molecular weight of the mature protein suggests carboxyl-terminal proc
                                                                                      A)Cross-references: EMBL: Z29076; NID: 9509266; PIDN: CAA82317.1; PID: 9509267
R; Seo, Y.S.; Lee, Y.H.; Pek, U.H.; Kang, H.
R; Seo, Y.S.; Microbiol. 24, 236-242, 1986
A; Title: Analysis on the nucleotide sequence of the signal region of Bacillus subitilis A; Reference number: 139803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation F;1-38/Domain: signal sequence #status predicted <SIG>F;39-508/Product: cellulase #status predicted <MAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
40.5%; Score 190; DB 2;
Best Local Similarity 40.7%; Pred. No. 4.2e-12;
Matches 35; Conservative 21; Mismatches 26;
                                                                                                                                                                                                                                                                                                                                            Status: preliminary, translated from GB/EMBL/DDBJ. Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :||||:
119 OIGCGNVTHKFVTLHKPKOGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63 AMGCGNIRASFGSVNPATPTADTYLO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Pathway: cellulose degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A, Accession: B26874
A, Molecule type: protein
A, Residues: 39-53 <ROB2>
                                                                                                                                                                                                                                                                                                         Accession: I39803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gene: bglC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names: endo-1,4-beta-glucanase
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate names
Capturate nam
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A, Wolecule type: DNA
A, Residues: 10-508 cMAC>
A; Residues: 10-508 cMAC>
A; Cross-rences: GB:X04689; NID:g39823; PIDN:CAA28392.1; PID:g39824
A; Experimental source: strain PAP115
A; Dindahl, V.; Aa, K.; Tronsmo, A.
Antonie Van Leeuwenhoek 66, 327-332, 1994
A; Title: Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus subtilis C
A; Accession: 140353
A; Accession: 140353
A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Experimental source: strain 168
Myanakay, R.M.; Lo, A.; Willick, Myllick, R.; Dove, M.; Moranelli, F.; Seli
Nucleic Acids Res. 14, 9159-9170; 1986
A; Title: Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.
A; Reference number: A26114; MUID:87066783; PMID:3024130
A; Accession: A26114
A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A,Description: Genes encoding beta glucan-hydrolyzing enzymes in Bacillus subtilis: cons
A,Reference number: S49103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Residues: 1-508 <KUN>
A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13696.1; PID:e1183471
                                                                                                                                                                                             63
                                                                                                                                          4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: EMBL:X67044; NID:g39776; PIDN:CAA47429.1; PID:g39777 R;Lindahl, V.; Aa, K. submitted to the EMBL Data Library, June 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Molecule type: DNA
A;Residues: 10-291, N', 293-508 <LINZ>
A;Cross-references: EMBL:X67044; NID:g39776; PIDN:CAA47429.1; PID:g39777
A;Experimental source: strain CK-2
R;Wolf, M.; Geczi, A.; Borriss, R.
            Pred. No. 1.4e-11;
4; Mismatches 33; Indels
43.5%; Preu. ....
                                                                                                                                                                                                                                                                                                                                                                                  1172 IGASNVTFNFVKLTSGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues: 10-291, 'N', 293-508 <LIN1>
                                                        37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Reference number: S24239
A;Accession: S24239
        Best Local Similarity
Matches 37; Conserv
                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                     ò
```

```
A;Experimental source: strain DLG
A;Note: the authors believe Met-1 and Met-2 may be alternate initiators
C;Comment: The low molecular weight of the mature protein suggests carboxyl-terminal proc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appescription: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel Apathway: cellulose degradation C;Reywords: alternative initiators; extracellular protein; glycosidase; hydrolase; polyse F;1-38/Domain: (or 2-38) signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2
                                                                                                    62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                                                                                                                                                                                                 cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain DLG)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Oct-1999
C;Accession: A26874; B26874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                               4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gapa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Reaidues: 1-508 <ROB1>
A;Cross-references: GB:M16185; NID:g143007; PIDN:AAA22496.1; PID:g143008
A;Experimental source: strain DLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4,
   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 508;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DIG.
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Accession: A268/4; proc).
R;Robson, L.M.; Chambliss, G.H.
A;Robson, L.M.; Chambliss, G.H.
A; Bacteriol. 169, 2017-2025, 1987
A;ritle: Endo-bera-1-4-glucanase gene of Bacillus subtilis |
A;Reference number: A26874; MUID:87194581; PMID:3106328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 40.5%; Score 190; DB 2; Best Local Similarity 43.0%; Pred. No. 4.2e-12; Matches 37; Conservative 17; Mismatches 28
```

```
Callulase (EC 3.2.1.-) precursor, thermoactive - Caldocellum saccharolyticum C.Species: Caldocellum saccharolyticum C.Species: Caldocellum saccharolyticum C.Species: Ladocellum saccharolyticum C.Accession: T17120; A43745

R.Te'o, V.S.; Saul, D.J.; Bergquist, P.L.
Appl. Microbiol. Biotechnol. 43, 291-296, 1995

A.Title: CelA, another gene coding for a multidomain cellulase from the extreme thermophical censor or T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Accession: T17120

A.Acce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-499 <NAK>
A; Residues: 1-499 <NAK>
A; Cross-references: GB:M28332; NID:g142670; PIDN:AAA223307.1; PID:g142671
A; Experimental source: strain IF03034
C; Function:
C; Function: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as capacity per period and period and period as capacity of period and period and period as capacity of period and period and period and period and capacity period and sequence #status predicted <SIG>
                                                                                             364 SGQIKVLYANKETNSTTNTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAIS-DW 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Species: Bacillus subtilis
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 15-Oct-1999
C;Accession: A27198
Evr. J. Biochem, A.; Uozumi, T.; Beppu, T.
Bur. J. Biochem, 164, 317-320, 1987
A;Reference number: A27198; MUID:87190397; PMID:3106035
A;Reference number: A27198; MUID:87190397; PMID:3106035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1516-1544,'A',1546-1742 <LUE>
A;Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72860.1; PID:g144293
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cellulase (EC 3.2.1.4) precursor - Bacillus subtilis (strain IFO3034) N.Alternate names: endo-1,4-beta-glucanase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels ~ 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1742;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .;
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 39.8%; Score 186.5; DB 2
Best Local Similarity 42.5%; Pred. No. 3.7e-11;
Matches 37; Conservative 18; Mismatches 31
                                                                                                                                                                                                                                                                          763 AQIGASNVTFKFVKLSSSVSGADYYLE 789
                                                                                                                                                                                                                         88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88
                                                                                                                                                                                                                         62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 AAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Gene: celA
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A43745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                요
                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                           Carboniase (EC 3.2.1.4) precursor - Bacillus subtilis (strain BSEG16)

N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus subtilis
C;Species: Bacillus subtilis
C;Species: Bacillus subtilis
C;Accession: JN0111

R;Park, S.H.; Kim, H.K.; Pack, M.Y.
Agric. Biol. Chem. 55, 441-448, 1991
A;Title: Characterization and structure of the cellulase gene of Bacillus subtilis BSEG1
A;Title: Characterization and structure of pmD:1368694
A;Reference number: JN0111; MUD:91299280; PMID:1368694
A;Reference number: JN0111
A;Molecule type: DNA
A;Residues: 1-499 cpAR>
A;Residues: 1-499 cpAR>
A;Residues: Translated the codon ATA for residue 102 as Tyr
C;Comment: The low molecular weight of the mature protein suggests carboxyl-terminal prc
C;Function:
A;Mosciption: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A;Bethway: cellulose degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide egradation
F;1-29/Domain: signal sequence #status predicted cSIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: nucleic acid
A; Residues: 1-1331 (GIB>
A; Residues: 1-1331 (GIB>
A; Residues: 1-1331 (GIB>
A; Residues: 1-1331 (GIB>
A; Crose-references: GB:L01257; NID:g144290; PIDN:AAA71887.1; PID:g144291
A; Crose-reference extracted from NCBI backbone (NCBIN:121576, NCBIP:121577)
B; Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
Appl. Environ. Microbiol. 57, 694-700, 1991
A; Title: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding A; Reference number: A43745; MUID:91247819; PMID:2039230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N;Alternate names: beta-mannanase
C;Species: Caldocellum saccharolyticum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A48954; B43745
R;Glbbs, M.D.; Saul, D.J.; Luthi, E.; Bergquist, P.L.
R;Glbbs, M.D.; Saul, D.S.; Sasse-1867, 1992
A;Title: The beta-mannanase from "Caldocellum saccharolyticum" is part of a multidomain
A;Reference number: A48954; MUID:93119139; PMID:1476429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         353 GISVQYRAGDGSMASNQIRPQLQIKNNGNTTVDLKDVTARYWYNAKNKGONV--DCDYAO 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     - Caldocellum saccharolyticum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Residues: 1-337, 'PPRQHORRQ' <LUE>
A; Cross-references: EMBL:M36063; NID:g144292; PIDN:AAA72861.1; PID:g144294
A; Orde: the authors translated the codon CAC for residue 262 as Glu
A; Note: this sequence has been revised in reference A49954
C; Keywords: glycosidase; hydrolase; polysaccharide degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 39.8%; Score 186.5; DB 2; Length 1331; Best Local Similarity 42.5%; Pred. No. 2.8e-11; Matches 37; Conservative 18; Mismatches 31; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ch 40.1%; Score 188; DB 2; 1 Similarity 40.0%; Pred. No. 6.7e-12; 34; Conservative 18; Mismatches 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mannan endo-1,4-beta-mannosidase (EC 3.2.1.78)
|||||: | :: |||||: QMGCGNLTHKFVTLHKPKQGADTYLE 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64 MGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 34; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: A48954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     411
```

g

ठ

g

ઠ

çej

4

Gaps

Ή,

62

g ò 셤

```
C;Species: Bacillus lautus
C;Date: 17-Peb-1994 #sequence_revision 17-Feb-1994 #text_change 15-Oct-1999
C;Accession: A41897; S27498
R;Hansen, C.K.; Diderichsen, B.; Jorgensen, P.L.
J. Bacteriol: 174, 3522-3531, 1992
A;Fitle: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-1,4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Accession: I40548
R;Han, S.J.; Yoo, Y.J.; Kang, H.S.
J. Biol. Chem. 270, 26012-26019, 1995
J. Biol. Chem. 270, 26012-26019, 1995
J. Ritle: Characterization of a bifunctional cellulase and its structural gene: the cel generaterization of a bifunctional PMID:7592793
                                                                                                                                                                                                                                              420 GQIKVLYANKETNSTTNTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAVS-DWA 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Species: Bacillus sp.
.Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Cross-references: EMBL: M76588; NID: g142661; PIDN: AAA22302.1; PID: g142662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: EMBL:U27084; NID:g857575; PIDN:AAC43478.1; PID:g857576
F;29-1039/Product: cellulase #status predicted <MAT> P;72-373/Domain: Streptomyces endo-1,4-beta-xylanase A homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38.4%; Score 180; DB 2; Length 486; 38.8%; Pred. No. 4.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: PL236
A;Note: sequence extracted from NCBI backbone (NCBIP:104604)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 38.5%; Score 180.5; DB 2; Best Local Similarity 43.9%; Pred. No. 1.1e-11; Matches 36; Conservative 18; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32;
                                                                                         DB 2;
                                                                                         38.9%; Score 182.5; DB 2 41.9%; Pred. No. 5.5e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18; Mismatches
                                                                                                                                                    18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 bifunctional cellulase precursor - Bacillus sp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cellulase homolog - Bacillus lautus (fragment)
                                                                                                                                                                                                                                                                                                                                                                      63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67 GNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         33; Conservative
                                                                                                                                                 36; Conservative
                                                                                      Query Match
Best Local Similarity
Matches 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 33; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A,Accession: A41897
A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-145 <HAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Residues: 1-486 < RES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: I40548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Gene: cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       exo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Accession: A43802
R;Saul, D.J.; Williams, L.C.; Grayling, R.A.; Chamley, L.W.; Love, D.R.; Bergquist, P.L.
Appl. Brvixon. Wicrobiol. 56, 3117-3124, 1990
A;Title: celB, a gene coding for a bifunctional cellulase from the extreme thermophile "A;Reference number: A43802; MUID:91136262; PMID:2126700
                                                                                                                                                                                                                                                                                                                                                                                                                                                 cellulase (EC 3.2.1.4) / cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91) - Caldocellum NiAlternate names: endo-1,4-beta-glucanase
NiAlternate names: endo-1,4-beta-glucanase
C.Species: Caldocellum saccharolyticum
C.Date: 30-Jan-1993 #sequence_revision 30-Sep-1993 #text_change 10-Jul-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL:X13602; NID:g40645; PIDN:CAA31936.1; PID:g40646
C;Genetics:
C;Genetics:
C;Genetics:
C;Genetics:
C;Function:
C;Function:
A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as (A;Pescription: hydrolysis of 1,4-beta-xylanase A homology C;Reywords: glycosidase; hydrolase; polysaccharide degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A homology C;Reywords: glycosidase; hydrolase; polysaccharide degradation
C;Fil-28/Domain: signal sequence #status predicted <SIG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as A;Pathway: cellulose degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A homology
C;Keywords: glycosidase; hydrolase; polyaexcharide degradation
F;20-320/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fitle: Nucleotide sequence of a gene from Caldocellum saccharolyticum encoding for Reference number: S02711; MUID:89098398; PMID:2789517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NiAlerance manes: endo-1.4-beta-glucance maccanain procession: NiAlerance manes: endo-1.4-beta-glucance manes: endo-1.4-beta-glucance manes: endo-1.4-beta-glucance manes: endo-1.4-beta-glucance manes: caldocallum saccharolyticum
C;Species: Caldocallum saccharolyticum
C;Species: Caldocallum saccharolyticum
C;Accession: S02711
R;Saul, D.J.; Williams, L.C.; Love, D.R.; Chamley, L.W.; Bergquist, P.L.
Nucleic Acids Res. 17, 439, 1989
A;Title: Nucleotide sequence of a gene from Caldocallum saccharolyticum encoding for A;Reference number: S02711; MUID:89098398; PMID:2789517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ij
                                                             3
                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Residues: 1-1039 <SAU>
Cross-references: EMBL:X13602; NID:g40645; PIDN:CAA31936.1; PID:g40646
                                                                4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2; Length 915;
      Length 499;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cellulase (EC 3.2.1.4) precursor - Caldocellum saccharolyticum
                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31; Indels
                                                             28;
39.7%; Score 186; DB 2;
40.7%; Pred. No. 1.1e-11;
iive 19; Mismatches 28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  38.9%; Score 182.5; DB 2
41.9%; Pred. No. 4.8e-11;
ive 18; Mismatches 31
                                                                                                                                                                                                                                                                                                  410 QIGCGNLTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OIGASNVTFKFVKLSSSVSGADYYLE 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   88
                                                                                                                                                                                                                                           63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63 AMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cross-references: EMBL:X13602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 41.9
                                                             Conservative
                        Local Similarity
hes 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Residues: 1-915 <SAU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Molecule type: DNA
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Function:
                                     Best Loc
Matches
```

염 ò 셤

6

7

Gaps

3;

99

63

Gaps

:||||: ::: ||||: LGCGNVSHTVVTLHKPKQGADTYLE 435

cellulase (EC 3.2.1.4) - Bacillus lautus

88

MGCGNIRASFGSVNPATPTADTYLQ

64

ሯ g

```
CipA protein - Clostridium thermocellum
NiAlternate names; probable cellulosome protein large chain SL
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: Clostridium thermocellum
CiSpecies: U.T.; Demain, A.L.
Submitted to the EMBL Data Library, January 1993
A;Reference number: S36859
A;Reference number: S36859
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1854 <GER>A;Residues: 1-1854 <GER>A;Residues: 1-1854 <GER
A;Cross-references: EMBL:LO8665
A;Cross-references: EMBL:LO8665
A;Cross-references: EMBL:LO8665
A;Cross-references: EMBL:LO8665
A;Title: Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal A;Reference number: S33527; MUID:93302508; PMID:8316083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1615,161-1854 <GE2>
A;Residues: 1-1615,161-1854 <GE2>
A;Cross-references: EMBL:L08665
R;Fujino, T.; Beguin, P.; Aubert, J.P.
FEMS Mitrobiol. Lett. 94, 165-170, 1992
A;Title: Cloning of a Clostridium thermocellum DNA fragment encoding polypeptides that b: A;Reference number: S25767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A,Accession: S28659
A,Molecule type: DDA Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1768,'R',1770-1854 ·
A,Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1768,'R',1770-1854 ·
A,Cross-references: EMBL:X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R.Fujino, T.; Beguin, P.; Aubert, J.P.
J. Bacteriol. 175, 1891-1899, 1993
A;Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulosom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Accession: $25767
A;Molecule type: DNA
A;Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1854 <FUJ>
A;Cross-references: EMBL:X67406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: EMBL:X67506; NID:9296879; PID:9296880; PIDN:CAA47840.1
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 35.2%; Score 165; DB 2; Length 1854; Best Local Similarity 36.1%; Pred. No. 6.9e-09; Matches 35; Conservative 23; Mismatches 29; Indels 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      424 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Reference number: Z18847; MUID: 93209931; PMID: 8458832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 WAAM-----GCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1821-1854 <FU2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;Beguin, P. submitted to the EMBL Data Library, August 1992 A;Reference number: S28659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: T18261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Gene: cipA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A;Pathway: cellulose degradation
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Accession: A47704; MUID:93171873; PMID:8436949

C;Accession: A47704; MUID:93171873; PMID:8436949
                                                                                                                                                                                                                                                                                                                                 C;Accession: B41897; S27499

R;Hansen, C.K.; Diderichsen, B.; Jorgensen, P.L.

J. Bacteriol. 174, 3522-3531, 1992

A;Title: celA from Bacillus lautus PL236 encodes a novel cellulose-binding endo-beta-1,4

A;Reference number: A41897; MUID:92276330; PMID:1592807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ٠.
س
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ϊ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              endoglucanase I (EC 3.2.1.-) CelI - Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
                                                                                                                                                                                                                                           N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus lautus
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Oct-1999
C;Accession: B41897; S27499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Cross-references: EMBL:M76588; NID:g142661; PIDN:AAA22303.1; PID:g142663
A;Experimental source: PL236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: GB:L04735; NID:g144807; PIDN:AA20892.1; PID:g144808
A;Note: sequence extracted from NCBI backbone (NCBIN:125637, NCBIP:125638)
C;Keywords: glycosidase; hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ٦;
```

Ouery Match 36.7%; Score 172; DB 2; Length 700; Best Local Similarity 44.9%; Pred. No. 4.5e-10; Matches 40; Conservative 18; Mismatches 27; Indels

| | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88

608

g

à

g

ঠ

A; Note: sequence extracted from NCBI backbone (NCBIP:104605)

A; Molecule type: DNA A; Residues: 1-700 <HAN>

Status: preliminary

, . , .

Search completed: May 11, 2004, 12:10:26 Job time: 4.1265 secs

61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88

RESULT

35.3%; Score 165.5; DB 2; Length 879; 33.0%; Pred. No. 2.7e-09; tive 26; Mismatches 32; Indels 1

Local Similarity 33.0% hes 29; Conservative

ઠે d ð g

Query Match Best Loca Matches

A; Molecule type: DNA; protein A; Residues: 1-879 <HAZ>

A; Status: preliminary

 $\boldsymbol{\dashv}$

7

```
GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 11:41:35 ; Search time 2.39913 Seconds

(without alignments)

1909.933 Million cell updates/sec
```

Run on:

May 11, 2004, 11:41:35; Search time 2.39913 Seconds (without alignments)

(without alignments)

1909.933 Million cell updates

Title:

US-09-917-376-5

Perfect score: 469
Sequence:

1 VSGGVKVQYKNNDSAPGDNQ......IRASFGSVNPATPTADTYLQ 88
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0 Maximum DB seq length: 200000000 Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Match Length DB ID Match Length DB ID 40.5 499 1 GUN1_BACSU P07981 Dacillus 40.1 499 1 GUN1_BACSU P07981 Dacillus 40.1 499 1 GUN1_BACSU P07981 Dacillus 40.1 499 1 GUN1_BACSU P07981 Dacillus 40.1 499 1 GUN1_BACSU P07981 Dacillus 73.8 1742 1 GUN1_CALSA P0.2533 caldocell 38.9 1742 1 GUN1_CALSA P0.2533 caldocell 38.9 1742 1 GUN1_CALSA P0.2534 caldocell 38.3 17.0 1 GUN1_CLOTTM P0.2534 clostrid 23.2 1853 1 GUN1_CLOTTM P0.2534 clostrid 23.2 1853 1 GUN1_CLOTTM P0.2534 clostrid 23.2 1853 1 GUN1_CLOTTM P0.2534 clostrid 23.1 GUN1_CLOTTM P0.2534 clostrid 23.1 GUN1_CLOTTM P0.2534 clostrid 23.2 1853 1 GUN1_CLOTTM P0.2534 clostrid 23.2 1853 1 GUN1_CLOTTM P0.2534 clostrid 23.2 1848 1 GUN1_CLOSR P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 GUNN_ERWCA P0.2639 erwinia 31.4 504 1 FLAI_THEW P0.2639 erwinia 22.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 FLAI_THEW P0.2639 erwinia 31.3 80.2 1 HISS_RHOSH P0.2639 FD.2639 erwinia 31.3 80.2 1 HISS_RHOSH P0.2639 FD.2639 FD.2639 Erwinia 31.3 80.2 1 HISS_RHOSH P0.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.2639 FD.26	Result		* Query			SUMMAKIES	
190 40.5 499 1 GUNL_BACSU P07983 190 40.5 499 1 GUNL_BACSU P07983 186.5 39.8 133.1 1 MANB_CALSA P22534 186.5 39.8 133.1 1 MANB_CALSA P22534 186.5 39.8 1742 1 GUNB_CALSA P22534 180.5 38.9 1039 1 GUNB_CALSA P22534 180.5 38.9 1039 1 GUNB_CALSA P22534 180.5 38.9 1039 1 GUNB_CALSA P22534 165.5 38.9 1039 1 GUNB_CALCT P22719 165.5 35.3 37.9 1 GUNB_CROT Q06851 149.5 31.9 504 1 GUNB_CROT Q10865 147.5 31.4 504 1 GUNB_CROT Q5336 147.5 31.4 504 1 GUNB_CROT Q5336 147.5 31.4 504 1 GUNB_CROT Q5336 147.5 31.4 504 1 GUNB_CROT </td <td>. o.</td> <td>Score</td> <td>Match</td> <td>Length</td> <td>BB</td> <td>qı</td> <td>Description</td>	. o.	Score	Match	Length	BB	qı	Description
190 40.5 499 1 GUNZ BACSU P10475 186.5 39.8 1742 1 GUNZ BACSU P22534 186.5 39.8 1742 1 GUNZ CALSA P22534 186.5 39.8 1742 1 GUNZ CALSA P22534 186.5 38.9 1039 1 GUNZ CALSA P22534 180.5 38.9 1039 1 GUNZ CALSA P22534 180.5 38.9 1039 1 GUNZ CALSA P22534 165.5 35.3 400 1 GUNZ CALOTM Q06851 165.5 31.9 444 1 GUNZ CLOTM Q06851 149.5 31.9 444 1 GUNZ CLOTM Q05934 147.5 31.4 504 1 GUNZ CLOSR P5090 147.5 31.4 504 1 GUNZ CLOSR P5395 147.5 31.4 504 1 GUNZ CLOSR P5305 141.5 30.2 1848 1 CHOZ CLOSR P5305 141.5 31.4 1 HV3 MOUSE P	ч	190		499	н		
186 40.1 499 1 GUN3 BACSU P23549 186.5 39.8 1331 1 MANB_CALSA P25533 186.5 39.8 1732 1 GUNA CALSA P25533 180.5 38.9 1039 1 GUNA CALSA P10474 180.5 38.9 1039 1 GUNA CALSA P259718 165.6 35.3 47 1 GUNA PABLA P29719 165.7 36.7 1 GUNA CLOTM 002934 165.8 35.2 1873 1 CIPA CLOTM 002934 149.5 31.9 444 1 GUNV ERWCA 002934 147.5 31.4 50.5 1 GUNV ERWCA 059394 147.5 31.4 50.4 1 GUNV ERWCA 059394 147.5 31.4 50.4 1 GUNV ERWCA 059394 147.5 31.4 1 HA3 MALO 059394 141.5 30.2 30.8 1 PUNZ CLOSR P50300 141.5 31.4 1 HA3 MALO 059316 071606 <	7	190		499	H	GUN2_BACSU	
186.5 39.8 1331 MANB CALSA P22534 186.5 39.8 1742 1 GUNA_CALSA P22534 186.5 39.8 1742 1 GUNA_CALSA P22534 180.5 38.9 1742 1 GUNA_CALSA P22534 180.5 38.5 145 1 CUNA_CALSA P22514 165.3 35.2 172 1 GUNA_CALCAM Q01864 165.3 35.2 772 1 CIPB_CLOTM Q01864 149.5 31.9 449 GUNV_ERWCA Q59394 147.5 31.4 504 1 GUNV_ERWCA Q59394 147.5 31.4 504 1 GUNV_ERWCA Q59394 147.5 31.4 504 1 GUNV_ERWCA Q59394 147.5 31.4 504 1 GUNV_ERWCA Q59395 147.5 31.4 504 1 GUNV_ERWCA Q59365 67.5 14.4 118 1 HV39_MOUSE P000SE 67.5 14.4 118 1 HV39_MOUSE	ო	188		499	~	GUN3_BACSU	
186.5 39.8 1742 I GUNA CALSA P22534 186.5 38.9 1039 I GUNB_CALSA P2234 180.5 38.9 1039 I GUNB_CALSA P2234 180.5 38.9 1039 I GUNB_CALSA P29718 165.5 38.3 48.9 I GUNB_CALCT Q02934 165.5 31.3 37.9 I GUNI_CATM Q06851 149.5 31.9 44.4 I GUNI_CATM Q06851 147.5 31.9 504 I GUNI_ERWCA Q47096 147.5 31.4 504 I GUNI_ERWCA Q53394 147.5 31.4 41 I HV3 M0USE P53658 67.5 14.4 118	4	186.5		1331	Н	MANB_CALSA	
182.5 38.9 1039 I GUNB CALSA P10414 180.5 38.9 1039 I GUNB CALSA P10414 180.5 38.5 145 I CYCRA-PAELA P29718 165.5 35.3 87.9 I GUNA CLOTM Q02934 165.5 35.2 187.2 I CIPA-CLOTM Q02934 149.5 31.9 44.4 I GUNV ERWCA Q67086 147.5 31.4 50.4 I GUNV ERWCA Q67939 147.5 31.4 50.4 I GUNZ CLOSR P50900 147.5 30.2 986 I GUNZ CLOSR P50900 67.5 14.4 118 I HV3 CLOSR P50900 67.5 14.4 118 I HV3 CLOSR P50900 67.5 14.4 118 I HV	ហ	186.5		1742	-	GUNA_CALSA	
180.5 38.5 145 1 YCEA PAELA PAELA P29718 165.35.3 36.3 1 GUNA_PAELA 0002918 165.35.2 1853 1 GUNI_CLOTM 0002851 149.5 31.9 44 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 504 1 GUNN_ERWCA 059395 147.5 31.4 118 1 HV39_MOUSE P53658 66.5 14.2 246 1 TPIA_MALDO 0954897 66.5 13.9 52.3 1 PUR9_MYCTU 0871553 64.5 13.6 1571 1 ATCS_YEAST P53696 63.5 13.5 247 1 FYA1_THEVO 0971563 63.1 13.4 1581 1 PPRB_HUMAN 099UPV9 63.1 13.4 1581 1 PPRB_HUMAN 099UPV9 62.1 13.2 1510 1 TCEN_PEREL P599156 62.1 13.1 16EN_PEREL P59ERE 61.5 13.1 16EN_PERER P59180 61.5 13.1 16EN_PERER P609915 61.5 13.1 16EN_PERER P609915 61.5 13.1 16EN_PERER P609915	9	182.5		1039	Н	GUNB_CALSA	
165.5 36.7 700 1 GUNA PABLA P29719 165.5 35.3 87.9 1 GUNA PABLA P29714 165.3 35.3 87.9 1 GUNI CLOTM Q00881 165.3 31.9 44.4 1 GUNI ERWCA Q05334 149.5 31.9 94.4 1 GUNI ERWCA Q59394 147.5 31.4 50.4 1 GUNI ERWCA Q47096 147.5 31.4 50.4 1 GUNI ERWCA Q59395 114 24.3 1848 1 GBPA_CLOCK Q47096 67.5 14.4 118 1 HV39 MOUSE P031609 67.5 14.4 118 1 HV39 MOUSE P031609 67.5 14.4 216 1 FLAI METVA P95116 66.5 13.9 52.3 1 PUR9 MYCTU Q95697 64.5 13.6 52.3 1 LIVD_THETV Q95697 64.5 13.6 52.1 1 LY3H_GCUN Q95719 65.1 13.1 14.3 188B_HUMAN Q910PV 65.1 13.3 26.9 1 FLAI MEN PPRB_HUMAN Q910PV 65.1 13.3 26.9 1 FLAI MEN PPRB_HUMAN Q910PV 65.1 13.3 26.9 1 FTNB MOUSE P05915 66.5 13.3 26.9 1 TYB_GCUN P05915 67.5 13.3 25.3 1 HORD MOUSE P05915 68.5 13.3 25.3 1 HORD MOUSE P05915 69.5 13.2 25.3 1 HORD MOUSE P05915 60.1 25.3 1 HORD MOUSE P05915 61.5 13.1 25.2 1 HORD MOUSE P05915 61.5 13.1 25.2 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD MOUSE P05915 61.5 13.1 316 1 HORD	7	180.5		145	Н	YCEA_PAELA	
165.5 35.3 879 I GUNI CLOTM 002934 165 35.2 37.2 1 CIPB_CLOTM 001866 149.5 31.9 44.4 1 GUNN ERWCA 069394 149.5 31.9 50.5 1 GUNV ERWCA 069394 147.5 31.4 50.4 1 GUNV ERWCA 069394 147.5 31.4 50.4 1 GUNV ERWCA 069394 141.5 30.2 986 1 GUNZ CLOSR P50900 141.5 30.2 986 1 GUNZ CLOSR P50900 67.5 14.4 118 1 HV39_MOUSE P60900 67.5 14.4 215 1 FLAI_METVA 091809 67.5 14.4 215 1 FLAI_METVA 091809 67.5 13.9 52.3 1 PUND_THETVA 091809 64.5 13.6 15.1 1 ATCS_TRAST P3266 64.5 13.5 24.7 1 FLAI_THETN 091804 63.5 13.5 24.7 1 FLAI_THEN	œ	172		700	Н	GUNA_PAELA	
165 35.2 772 1 CIPB CLOTM 166 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165 165	σ	165.5		879	H	GUNI_CLOTM	
165 35.2 1853 CIPA_CLOTM 169 149.5 31.9 444 10 CINN_ERWCA 149.5 31.9 444 10 CINN_ERWCA 149.5 31.9 444 10 CINN_ERWCA 147.5 31.4 504 10 CUNN_ERWCA 147.5 31.4 504 10 CUNN_ERWCA 147.5 31.4 504 10 CUNN_ERWCA 147.5 14.4 118 10 CLOSR P38058 67.5 14.4 215 17 FLAI METVA P95316 66.5 14.2 246 1 FLAI METVA P95316 66.5 13.9 52.3 1 FLAI METVA P95316 66.5 13.9 52.3 1 FLAI METVA P95316 66.5 13.9 52.3 1 FLAI METVA P971553 64.5 13.6 15.7 1 ATCS_YEAST P971553 66.5 13.5 24.7 1 FLAI THEVO P57179 67.5 13.4 55.3 1 FLAI THEVO P57179 67.5 13.4 55.3 1 FLAI THEVO P57179 67.5 13.3 26.3 1 FLAI MAN Q910PV9 63 13.4 1581 1 PPRB_HUMAN Q910PV9 63 13.4 1581 1 PPRB_HUMAN Q910PV9 63 13.2 1 CEN PERFL P988 P988 P988 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188 P188	10	165		772	~	CIPB CLOTM	
149.5 31.9 444 1 GUNN ERWCA Q59394 149.5 31.9 505 1 GUNV ERWCA Q47096 147.5 31.4 504 1 GUNV ERWCA Q59395 141.5 30.2 986 1 GUNZ CLOSR P50900 141.5 30.2 986 1 GUNZ CLOSR P50900 67.5 14.4 118 1 HV39_MOUSE P01809 67.5 14.4 215 1 FLAI_METVA P01809 65 13.9 5.23 1 PUND_THETVA Q9458316 64.5 13.8 5.56 1 LUVD_THETVA Q84697 64.5 13.9 5.23 1 PUND_THETVA Q84719 64.5 13.6 15.1 1 ATCS_YEAST P55316 64.5 13.5 24.7 1 FLAI_THETN Q94749 63.5 13.5 24.7 1 FLAI_THEVA Q94746 63.5 13.5 46.5 1 TY3H_GUMA Q940PV9 62.5 13.4 1581 1 PRB_HUMA	11	165		1853	Н	CIPA_CLOTM	
149.5 31.9 505 1 GUNV ERWCA Q47096 147.5 31.4 504 1 GUNW_ERWCA Q59395 147.5 31.4 504 1 GUNW_ERWCA Q59395 141.5 30.2 986 1 GUNZ_CLOSR P50900 67.5 14.4 118 1 KPA_CLOCL P38058 67.5 14.4 218 1 KPA_CLOCL P38058 67.5 14.4 218 1 KPA_CLOCL P538058 64.5 14.2 246 1 FPIA_MALDO Q9f897 64.5 13.9 55.3 1 LVD_THETN Q9f897 64.3 13.6 1571 1 ATC_YEAST P87660 63.5 13.5 247 1 FALA_THEVO O17446 63.1 13.4 1581 1 PPRB HUMAN Q9upv9 62.5 13.3 80.2 1 EXD_THIME P33693 62.5 13.3 80.2 1 EXD_WERF Q9upv9 62.5 13.2 95.0 1 STLZ_MOUSE	12	149.5		444	Н	GUNN_ERWCA	
147.5 31.4 504 1 GUNW ERWCA Q59395 147.5 31.4 914 1 GUNZ_CLOSR P50900 141.5 30.2 986 1 GUNZ_CLOSR P50900 67.5 14.4 118 1 HV39_MOUSE P38058 67.5 14.4 215 1 FLA1 METVA P91809 66.5 14.2 246 1 FLA1 METVA P9516 65 13.9 52.3 1 PUR9 MYCTU Q945G7 64.5 13.6 55.3 1 ILVD THETN Q945G7 64.5 13.5 247 1 FLA1 THEVO Q945G7 63.5 13.5 247 1 FLA1 THEVO Q94D49 63.5 13.5 247 1 FLA1 THEVO Q194A6 63.5 13.4 1581 1 PPRB HUMAN Q94D49 62.5 13.3 269 1 FLAL MOUSE Q95D50 62.5 13.2 950 1 FLAL MOUSE Q99D50 62.13.2 12.0 1 FLAL MOUSE Q99D50	13	149.5		505	Н	GUNV ERWCA	
147.5 31.4 914 1 GUXZ_CLOSR P50900 141.5 30.2 986 1 GUNZ_CLOSR P50900 141.5 30.2 986 1 GUNZ_CLOCL P3655 67.5 14.4 118 1 HV39_MOUSE P01809 66.5 14.4 215 1 FLAI_METVA P95316 66.5 13.9 523 1 PURD_MALDO Q9458316 64.5 13.8 556 1 LUVD_THETN Q87159 63.5 13.1 ATCS_YEAST P71553 63.5 13.5 247 1 FLAI_THETN Q87419 63.5 13.5 247 1 FLAI_THEN P57119 63.5 13.5 247 1 FLAI_THEN P57719 63.5 13.5 247 1 FLAI_THEN Q940Y9 63.5 13.5 465 1 TY3H_GCHMA Q940Y9 63.13.4 1581 1 PRB_HUMAN Q15648 62.5 13.3 269 1 SEXOK_RHIMB P3369	14	147.5	31.4		М	GUNW_ERWCA	
141.5 30.2 986 1 GUNZ CLOSR P235659 114 24.3 1848 1 HV39 P00SE F78058 67.5 14.4 215 1 FLAI_METVA P95316 66.5 14.2 246 1 TP1A MALDO Q9f897 64.5 13.9 52.3 1 PUND_THETV P71553 64.5 13.6 1571 1 ATC5 YRAST P71553 63.5 13.6 1571 1 ATC5 YRAST P71553 63.5 13.4 1581 1 PALA_THEVO P71553 63.1 13.4 953 1 O106_HUMAN Q9upv9 62.5 13.3 26 1 XY3H SCHMA Q15648 62.5 13.3 80.2 1 PPRB HUMAN Q9upv9 62.5 13.3 80.2 1 PYNB MOUSE P29359 62.5 13.2 95.0 1 STL2 MOUSE Q99H50 62.13.2 13.1 1 ISCR_PSEFL P099H50 61.5 13.1 1 ISCR_PSEFL P099H50 </td <td>15</td> <td>147.5</td> <td>31.4</td> <td></td> <td>н</td> <td>GUX2_CLOSR</td> <td></td>	15	147.5	31.4		н	GUX2_CLOSR	
114 24.3 1848 1 CBPA CLOCL P38058 67.5 14.4 118 1 HV39_MOUSE P01809 67.5 14.4 218 1 FLA1 METVA P01809 66.5 14.2 246 1 FPIA MALDO Q9£897 64 13.6 15.7 1 PUR9 MYCTU Q9£897 64 13.6 1571 1 ATCS_YEAST Q8745 63.5 13.5 247 1 FAL1 THEVO Q9745 63.5 13.5 247 1 FAL1 THEVO Q1446 63 13.4 1581 1 PPRB HUMAN Q9upv9 62.5 13.3 262 1 EXOK_HHIME P3369 62.5 13.2 950 1 PTNB MOUSE Q99n50 62 13.2 950 1 STLZ MOUSE Q99n50 61.5 13.1 125 1 HSG RHOSH P59915 61.5 13.1 1 HSG RHOSH P59915 61.5 13.1 1 HSG RHOSH P59915 61.5 <td>16</td> <td>141.5</td> <td>30.2</td> <td></td> <td>Н</td> <td>GUNZ_CLOSR</td> <td></td>	16	141.5	30.2		Н	GUNZ_CLOSR	
67.5 14.4 118 1 HV39 MOUSE P01609 67.5 14.4 215 1 FLAI_METVA P95116 96.5 14.2 246 1 TPLA MALDO Q95 FS316 64.5 13.9 523 1 PUR9 MYCTU Q8745 64.5 13.6 15.7 1 LIVD_THETN Q8745 63.5 13.5 24.7 1 FLAI_THEVO P57119 63.5 13.5 24.7 1 FLAI_THEVO P57119 63.5 13.5 24.7 1 FLAI_THEVO P57119 62.5 13.3 26.9 1 O106 HUMAN Q99UPV9 63 13.4 1581 1 PPRB HUMAN Q99UPV9 62 13.2 950 1 PTNB MOUSE P5955 62 13.2 950 1 STLZ_MOUSE P5955 62 13.2 250 1 STLZ_MOUSE P5955 65 13.2 12.0 1 CEN PSEFL P09815 61.5 13.1 252 1 HISG RHOSH P38108	17	114	24.3	1848	Н	CBPA_CLOCL	
67.5 14.4 215 1 FLA1 METVA P95316 66.5 14.2 246 1 P1A_MALDO Q6587 66.5 13.9 523 1 PUR9 MYCTU Q65897 64.5 13.8 556 1 ILVD_THETN Q8rdj9 64.5 13.6 1571 1 FLA1_THEVO Q6759 63.5 13.5 24.7 1 FLA1_THEVO Q7466 63.5 13.4 95.3 1 O10_HUMAN Q9upv9 63 13.4 1581 1 PRRB HUMAN Q9upv9 62.5 13.3 26.9 1 EXOK RHIME P29359 62.5 13.3 26.9 1 EXOK RHIME P29359 62.5 13.2 950 1 STL2 MOUSE Q99150 62.5 13.2 1 ILSEN_PSEFL PROSE P29315 61.5 13.1 25.2 1 HISG RHOSH P3108 13.8 10.1 1 ILSEN_PSEFL P09815 61.5 13.1 316 1 MUCB_PSEAE P38108	18	67.5	14.4	118	н	HV39_MOUSE	_
66.5 14.2 246 1 TP1A MALDO Q95897 64.5 13.8 556 1 ILVD THETN Q8471 64. 13.6 1571 1 ATC_YEAST Q8460 63.5 13.5 247 1 FLAL_THEVO P57719 63.5 13.5 46.5 1 TY3H SCHMA Q910PV9 63.1 13.4 953 1 O106_HUMAN Q910PV9 62.5 13.3 26.9 1 EXOK_WHIME P33693 62.5 13.3 802 1 EXOK_WHIME P33693 62.5 13.2 950 1 STLZ MOUSE Q99150 63. 13.4 1581 1 CEN_FESFL P988 62.1 13.2 950 1 STLZ MOUSE P93955 64.5 13.2 110 1 ICEN_FESFL P99815 65.5 13.1 13.1 110 INCE_BEARE P38108	13	67.5	14.4		-	FLA1_METVA	
65 13.9 523 1 PUR9 MYCTU P71553 64.5 13.6 556 1 LLVD_THETN Q8rdj9 65.5 13.5 247 1 FLA1_THEVO P57159 63.5 13.5 247 1 FLA1_THEVO P57119 63.5 13.5 247 1 FLA1_THEVO P57119 62.5 13.3 269 1 PRB_HUMAN Q9upv9 63 13.4 1581 1 PPRB_HUMAN Q15648 62.5 13.3 269 1 PRRB_HUMAN Q15648 62.5 13.3 269 1 PTNB_MOUSE P59159 62 13.2 12.0 1 TCM_PSEFL P09159 63 13.2 12.0 1 LCM_PSEFL P0915 61.5 13.1 252 1 HISG_ROSH P38108 F38108	20	66.5	14.2		H	TP1A_MALDO	
64.5 13.8 556 1 ILVD THETN Q8rdj9 64.13.6 13.7 1 ATCS_YEAST P3266 63.5 13.5 247 1 FLA1 THEVO P57119 63.5 13.4 95.3 1 0106 HUMAN Q15466 62.5 13.4 1581 1 PARB_HUMAN Q15648 62.5 13.3 269 1 EXOK RHIME P29352 62.13.2 950 1 STL2 MOUSE Q99150 62.13.2 13.1 1 ICEN_PSFFL P09815 61.5 13.1 316 1 MUCB_PSEAE P38109	21	65	13.9		ᆏ	PUR9 MYCTU	
64 13.6 1571 1 ATC5 YEAST P32660 63.5 13.5 247 1 FAIA_THEVO P57719 63.5 13.4 4 65 1 YY3H SCHMA O17446 62.5 13.3 2.69 1 EXOK_HIME P33693 62.5 13.3 2.69 1 EXOK_HIME P33693 62.5 13.3 902 1 PYNB MOUSE P29352 62 13.2 950 1 STL2 MOUSE Q99150 62 13.2 13.1 1 1 CEN_PSEFL P09815 61.5 13.1 316 1 MUCB_PSEAE P38108	22	64.5	13.8		H	ILVD_THETN	_
63.5 13.5 247 1 FLA1 THEVO P57719 63.5 13.5 46.5 1 TY3H_GCHMA 017446 63.1 13.4 953 1 0106 HUMAN Q9bupv9 63 13.4 1581 1 PPRB HUMAN Q15648 62.5 13.3 269 1 EXOK_RHIME P33693 62 13.2 13.1 950 1 STL2_MOUSE P29352 62 13.2 1210 1 TCRN PSEFL P09815 61.5 13.1 252 1 HISG RHOSH P38108	23	64	13.6		H	ATC5_YEAST	
63.5 13.5 465 1 TY3H_SCHMA O17446 63 13.4 953 1 0.106_HUMAN Q9Lupv9 63 13.4 1581 1 PPRB_HUMAN Q15648 62.5 13.3 269 1 EXOK RHIME P33693 62.5 13.2 950 1 STL2_MOUSE P29352 62 13.2 950 1 STL2_MOUSE Q99n50 62 13.2 13.1 1 ICEN_PSFL P0915 61.5 13.1 316 1 MUCB_PSEAE P38108	24	3	13.5		Н	FLA1 THEVO	P57719 thermoplasm
63 13.4 953 1 0106_HUMAN Q9upv9 63 13.4 1581 1 PRB_HUMAN Q16468 62.5 13.3 269 1 EXOK RHIME P33693 62.5 13.2 950 1 STL2 MOUSE C2 13.2 950 1 STL2 MOUSE C2 13.2 13.1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	25	3	13.5		Н	TY3H SCHMA	O17446 schistosoma
63 13.4 1581 1 PPRB HUMAN Q15648 62.5 13.3 269 1 EXOK_HIME P33693 62.5 13.2 802 1 PTNB MOUSE P29352 62 13.2 13.1 1 1 TCRN PSBFL PSPR P99155 61.5 13.1 252 1 HISÉ RHOSH P38108 61.5 13.1 316 1 MUCB_PSRAE P38108	56	63	13.4		Н	O106_HUMAN	
62.5 13.3 269 1 EXOK_RHIME P33693 62.5 13.3 802 1 PTNB MOUSE P29352 62 13.2 950 1 STLZ_MOUSE Q99150 62 13.2 12.10 1 ICEN PSEFL P09915 61.5 13.1 252 1 HIS6_RHOSH P56937 61.5 13.1 316 1 MUCB_PSEAE P38108	27	63		1581	~	PPRB_HUMAN	
62.5 13.3 802 1 PTN8 MOUSE P29352 62 13.2 950 1 STL2 MOUSE Q99n50 62 13.2 12.10 1 ICEN_PSFFL P0915 61.5 13.1 252 1 HIS6 RHOSH P50931 61.5 13.1 316 1 MUCB_PSEAE P38108	28	N		269	н	EXOK_RHIME	
62 13.2 950 1 STL2 MOUSE Q99H50 62 13.2 1210 1 ICEN_ESFFL P09BH5 61.5 13.1 252 1 HIS6 RHOSH P50937 61.5 13.1 316 1 MUCB_PSEAE P38H08	29	~	•	802	ч	PTN8_MOUSE	
62 13.2 1210 1 ICEN_PSEFL P09815 61.5 13.1 252 1 HIS6_RHOSH P50937 61.5 13.1 316 1 MUCB_PSEAE P38108	30	62	•	950	Н	STL2 MOUSE	
61.5 13.1 252 1 HIS6_RHOSH P50937 61.5 13.1 316 1 MUCB_PSEAE P38108	31	62	•	1210	Н	ICEN_PSEFL	
61.5 13.1 316 1 MUCB_PSEAE P38108	32	-	•	252	-		_
	33	\vdash	•	316	Н	ıı	P38108 pseudomonas

P14428 mus musculu P03991 mus musculu	Q9a510 caulobacter P32138 escherichia P06329 mus musculu	Q8fcs8 escherichia P46849 escherichia	Q45733 bacillus th P18891 aleuria aur	P34446 caenorhabdi P54507 bacillus su
HA1Q_MOUSE HA1W_MOUSE	GATA_CAUCR YIHQ_ECOLI HV50_MOUSE	RTCA_ECOL6 RTCA_ECOL1 PTCA_ECOL1	C9CA_BACTO	PAT2_CAEEL COTN_BACSU
328 368	489 678 120	3338	342 1157 312	1226 261
13.1	13.1 13.1	0.0.0	12.9	12.8
61.5	61.5 61.5 61	555	60.5 60.5	60 59.5
34 35	36 37 38	6 4 4 6 0 t	4 4 4 1 2 6	44 5

ALIGNMENTS

```
유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RA Kunst F., Ogsawara N., Mosser I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Brouillet S., Bruschie L., Brans A., Braun M., Briggell S.C., Bron S.,
Brouillet S., Bruschi C.W., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Brouillet S., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galistia A., Galleron N.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
Aoris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Murita K., Lopidus A., Lardinois S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Lardinois S., Lauber J., Lazarevic V.,
RA Medina N., Mellado R.P., Mixuno M., Moestl D., Nekai S., Noback M.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Persecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Saddie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Schleich J., Seror S.J., Serror S.J., Serror P., Shin B.S.,
Soldo B.,
Solokin A., Tacconi E., Takahashi H., Takemaru K.,
Lindahl V., Aa K., Tronsmo A.;
"Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rose M., Entian K.D.; "New genes in the 170 degrees region of the Bacillus subtilis genome encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino
                                                                                                                                                                                                                                                                                                                                                 Moranelli F., Seligy V.; "Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene."; Nucleic Acids Res. 14:9159-9170(1986).
                                                                                                                                                            01-JUL-1989 (Rel. 11, Created)
01-JUL-1989 (Rel. 11, Last sequence update)
11-0CT-2003 (Rel. 42, Last annotation update)
Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
(Carboxymethyl-cellulase) (CMCASE) (Cellulase).
                                                                                                                                                                                                                                                                                                                                       Mackay R.M., Lo A., Willick G., Zuker M., Baird S., Dove M.,
                                                                                                                                                                                                                                                       Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus
                                                                                                                                    499 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Antonie Van Leeuwenhoek 66:327-332(1994).
                                                        410 QMGCGNLTHKFVTLHKPKQGADTYLE 435
                                         63 AMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                        MEDLINE=87066783; PubMed=3024130;
                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=95225656; PubMed=7710280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=97124194; PubMed=8969507;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=98044033; PubMed=9384377;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Microbiology 142:3097-3101(1996).
                                                                                                                                                                                                                             OR GLD OR EGLS OR BSU18130
                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             transporter.";
                                                                                                                                                                                                                                          Bacillus subtilis.
                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. STRAIN=PAP115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                  NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        subtilis CK-2."
                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=CK-2
                                                                                                                                    GUN2 BACSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=168;
                                                                                                          RESULT 2
GUNZ_BACSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              acid
                                                                                                                                   g
                                        ò
                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                62
Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambutt R., Wedler B., Wedler H., Weitzenegger T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumstein B., Yoshikawa H., Danchin A.;
"The complete genome sequence of the Gram-positive bacterium Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWF-TRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Antonie Van Leeuwenhoek 66:319-326(1994).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=95225655; PubMed=7710279;
Aa K., Flengsrud R., Lindahl V., Tronsmo A.;
Characterization of production and enzyme properties of an endo-beta-1,4-glucanase from Bacillus subtilis CK-2 isolated from compost soil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENDOGLUCANASE.
PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARIT
S -> N (IN REF. 2).
W; 8F735FF711B3BAE2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40.5%; Score 190; DB 1; Length 499; 40.7%; Pred. No. 1.3e-13; ive 21; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSP, 085465; 1A3H.
Subtlinst; BG10477; bg1C.
InterPro; 1PR01956; CB1.
InterPro; IPR01956; CB1.
InterPro; IPR01956; CB1.
InterPro; IPR01956; CB1.
InterPro; IPR01956; CB2.
InterPro; IPR01956; CB2.
InterPro; IPR01959; CB2.
InterPro; IPR01959; CB2.
InterPro; IPR01959; CB2.
InterPro; IPR01959; CB2.
InterPro; IPR01959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
INTERPPRO; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
InterPro; IPR05959; CB2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         *||||: | :: | |||: 410 QIGCGNVTHKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CAA47429.1; -. CAA97610.1; ALT INIT. CAB13696.1; ALT INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GUN3 BACSU STANDARD;
P23549;
01-NOV-1991 (Rel. 20, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55287 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, Z29076; CAA82317.1; -.
EMBL, X04689; CAA2392.1; -.
EMBL, K67044; CA47429.1; -.
EMBL, Z73234; CAA97610.1; ALT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 40.7% tes 35; Conservative
                                                                                                                                                                                                                                                                                                                     Nature 390:249-256(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29
499
169
257
499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; Z99113; CAB136
PIR; G69593; G69593.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        257
350
283
499 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 30-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hydrolases)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=CK-2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
GUN3_BACSU
ID GUN3_BA
AC P23549;
DT 01-NOV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
```

m

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 42.5*
Matches 37; Conservative
  Caldicellulosiruptor.
NCBL_TaxID=44001;
                                       FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
CONFLICT
                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
    ð
                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the buspean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1991 (Rel. 19, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Beta-mannanase/endoglucanase A precursor [Includes: Mannan endo-1,4-beta-mannosidase A (EC 3.2.1.78) (Beta-mannanase) (Endo-1,4-MANA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                     subtilis BSE616.";
Agric. Biol. Chem. 55:441-448(1991).
-!- CAPALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                               Park S.H., Kim H.K., Pack M.Y., "Characterization and structure of the cellulase gene of Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENDOGLUCANASE.
PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
CELLULOSE-BINDING (BY SIMILARITY).
(Rel. 20, Last sequence update)
(Rel. 41, Last annotation update)
se precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40.1%; Score 188; DB 1; Length 499; 40.0%; Pred. No. 2.2e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31; Indels
                                                                          Firmicutes; Bacillales; Bacillaceae; Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2E821E3D8BBACA04 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProDom; PD001947; CBD 3; 1.
PROSITE; PS0659; GLYGOSYL HYDROL F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
                           Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4 (Carboxymethyl-cellulase) (CMCASE) (Cellulase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40.0%; Pred. No. 18: Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1331 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :|||||:
LGCGNVTYKFVTLHKPKQGADTYLE 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                           HSSP, 085465; IA3H.
InterPro, IPR001956; CBD 3.
InterPro, IPR008965; Cellul bind.
InterPro, IPR001547; Glyco_hydro_5.
                                                                                                                         STRAIN=BSE616;
MEDLINE=91299280; PubMed=1368694;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        257 257 M
350 499 CI
499 AA; 55169 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00942; CBM 3; 1. Pfam; PF00150; cellulase; 1.
                                                                                                                                                                                                                                                                                                                                                     EMBL; D01057; BAA00859.1; -. PIR; JN0111; JN0111.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                              SEQUENCE FROM N.A.
                                                             Bacillus subtilis.
                                                                                      NCBI_TaxID=1423;
                                                                                                                                                                                                                                  hydrolases).
    01-NOV-1991
28-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MANB CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 fatches
   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OS GE DE COS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Celsius.
SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO CELLULASE FAMILY
A (FAMILY 5 OF GLYCOSYL HYDROLASES).
SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO CELLULASE FAMILY
J (FAMILY 44 OF GLYCOSYL HYDROLASES).
                                                                                                                                                galactoglucomannans.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- MISCELLANEOUS: This enzyme is most active at pH 6 and 80 degrees
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Gaps
MEDLINE=93119139; PubMed=1476429;
Gibbs M.D., Saul D.J., Luthi E., Bergquist P.L.;
"The beta-mannanase from 'Caldocellum saccharolyticum' is part of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRO/SER/THR-RICH (PT BOX).
CATALYTIC (ENDOGLUCANASE ACTIVITY)
PROTON DONOR (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 186.5; DB 1; Length 1331;
Pred. No. 1e-12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CATALYTIC (MANNANASE ACTIVITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRO/SER/THR-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
SUBSTRATE-RICH (PT BOX).
SUBSTRATE-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        340 346 TPTPTPT -> RQHQHRQ (IN REF. 2)
1331 AA; 146892 MW; FFBCA51BB8D8F0E0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                               linkages in mannans, galactomannans, glucomannans, and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; L01257; AAA71887.1; -.

EMBL; M36063; AAA72861.1; -.

FIR; A48954; A48954.

HSSP; Q06651; INBC.

InterPro; IPR001956; CBD 3.

InterPro; IPR0018965; Calful bind.

InterPro; IPR001847; CBM 3, 2.

Pfam; PF00150; calfulase; 1.

Probom; PF001697; CBD 3; 2.

Probom; PF001695; GALVGOSYL HYDROL F5; 1.

Probom; PF001695; GLYGOSYL HYDROL F5; 1.

Hydrolase; Glycosidase; Cellulose degradation; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BETA-MANNANASE/ENDOGLUCANASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -> P (IN REF. 2
                                                                                                        Appl. Environ. Microbiol. 58:3864-3867(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        39.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Multifunctional enzyme
SIGNAL 1 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    361
518
564
720
780
1331
162
257
257
338
                                                                                multidomain enzyme."
```

```
ProDom; PD011903;
                                                                                                                                                                                                                                    ACT_SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                              N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                         SITE
                                                                                                                                                                                                                                                                                                 Query Match
Best Local
                                                                                                           DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P10474;
                                                                 SIGNAL
                                                                                              DOMAIN
                                                                                                                                                         DOMAIN
                                                                                                                                                                           DOMAIN
                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CALSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUINB
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY BE TEMALLY 9 OF CHYCOSYL HYDROLASES).
-!- SIMILARITY: IN THE C-TEMAINAL SECTION; BELONGS TO CELLULASE FAMILY L (FAMILY 48 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                              Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C-TERMINAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
PTM: The linker region (also termed "hinge") may be a potential
site for proteolysis:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1516-1742 FROM N.A.

MEDLINE=91247819; PubMed=2039230;

Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;

Leuthing, sequence analysis, and expression in Escherichia coli of gene coding for a beta-mannanase from the extremely thermophilic bacterium 'Caldocellum saccharolyticum'.";

Appl. Environ. Microbiol. 57:694-700(1991).

-!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN ENDOGUNCANASE ACTIVITY ON CARBOXYMETHYLCELLULOSE. THE C-TERMINA DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
                                                                                                                                                                                                    28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase A)
                                                                                                                                                                                                                                                                                                                                                               Te'O V.S., Saul D.J., Bergquist P.L.; "cela, another gene coding for a multidomain cellulase from the extreme thermophile Caldocellum saccharolyticum."; Appl. Microbiol. Biotechnol. 43:291-296(1995).
                                                                                                                                                         PRT; 1742 AA
                                                                           AQIGASNVTFKFVKLSSSVSGADYYLE 449
                                              62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; 1,12742; AAA91086.1; -.
EMBL; M36063; AAA72860.1; -.
EMBL; L01257; -; NOT_ANNOTATED_CDS.
PIR; T17120; T17120.
HSSP; P2621; 1TF4.
InterPro; IPR001956; CBD 3.
InterPro; IPR009965; Cellul bind.
InterPro; IPR009965; Cellul bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001701; Glyco_hydro_9.
InterPro; IPR008928; Glyco_trans_6hp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00942; CBM 3; 3. — Pfam; PF002011; 01yGo hydro 48; 1. Pfam; PF00759; 01yCo hydro 9; 1. PRINTS; PR00844; GLHYDRLASE48.
                                                                                                                                                                                      (Rel. 19, Created)
(Rel. 34, Last seq
                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=44001
                                                                                                                                                                                       01-AUG-1991
                                                                                                                                                         CALSA
                                                                           423
                                                                                                                                                                       P22534;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    +
                                                                                                                                                       유
                                            ò
                                                                         g
```

ProDom; PD001947; CBD_3; 2.

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as a since content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                  704 SGOIKVLYANKETNSTTNTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAOSAIS-DW 762
                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUL-1989 (Rel. 11, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase/exoglucanase B precursor [Includes: Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
(Cellobiohydrolase); Exoglucanase (EC 3.2.1.91) (Exocellobiohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                   SGGVKVOYKNNDSAPGDNOIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                    1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleic Acids Res. 17:439-439(1989).
-!- FUNCTION: THIS PROTEIN IS MADE UP OF TWO DOMAINS: THE N-TERMINAL DOMAIN HAS EXOGLUCANASE ACTIVITY WHILE THE C-TERMINAL DOMAIN IS AN ENDOGLUCANASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Saul D.J., Williams L.C., Love D.R., Chamley I.W., Bergquist P.I., "Nucleotide sequence of a gene from Caldocellum saccharolyticum encoding for exocellulase and endocellulase activity.";
                                                                                                                                                   CELULOSE-BINDING (BY SIMILARITY).
LINKER ("HINGE") (PRO-THR BOX).
CATALYTIC 2.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
T -> A (IN REF. 2).
                                                                    POTENTIAL.

ENDOGLICANASE A.

CATALYTIC 1.

LINKER ("HINGE") (PRO-THR BOX).

CELLULOSE-BINDING (BY SINTLARITY).

"""" ("HINGE") (PRO-THR BOX).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
Glyco hydro 48; 1.
GLYCOSYL HYDROL F9 1; 1.
GLYCOSYL HYDROL F9 2; 1.
GLYCOSYL HYDROL F9 2; 1.
                                                                                                                                                                                                                                                                                                                  Length 1742;
                                                                                                                                                                                                                                                                  T -> A (IN REF. 2).
MW; 3F0699A2123EED07 CRC64;
                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                                                                                                 31;
                                                                                                                                                                                                                                                                                                                                     .4e-12
                                                                                                                                                                                                                                                                                                                                 ; Pred. No. 1.4e-
18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1039 AA
                                                                                                                                                                                                                                                                                                                    Score 186.5;
Pred. No. 1.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=89098398; PubMed=2789517;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (1,4-beta-cellobiohydrolase)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1989 (Rel. 11, Created)
                                                                                                                                                                                                                                        443
1545
193696 M
                                                                                                                                                                                                                                                                                                                  39.8%;
                                                                                                                                                                                                                                                                                                                                                 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                   Cellulose degradation;
                                                                                                                                                                1060
1112
1742
                                                                                                                                                                                                                                                                                   AA;
                                                                                                                                                                                                                                                                                                                                 Similarity
                 PROSITE; PS00592;
PROSITE; PS00698;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=44001;
                                                                                                                                                                                                                                                    443
1545
1742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CALSA
```

S

us-09-917-376-5.rsp

```
PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUNA PAE
P29719:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
GUNA PAELA
  RARRE SE S
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      420 GQIKVLYANKETNSTTNTIRPWLKVVNSGSSSIDLSRVTIRYWYTVDGERAQSAVS-DWA 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
16-COT-2001 (Rel. 40, Last annotation update)
17-COT-2001 (Rel. 10 protein in celA 5'region (Fragment).
18-Paenibacillus lautus (Bacillus lautus)
18-Paeria: Firmicutes; Bacillales; Paenibacillas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hansen C.K., Joergensen P.L., Diderichsen B.; "celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENDOGLUCANASE/EXOGLUCANASE B.
THR/PRO-RICH, TANDEM REPEATS OF T-P.
THR/BCD-BINDING (BY SIMILARITY).
THR/PRO-RICH, TANDEM REPEATS OF T-P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38.9%; Score 182.5; DB 1; Length 1039; ilarity 41.9%; Pred. No. 2.1e-12; Conservative 18; Mismatches 31; Indels 1;
                                                                                                    HSSP, OGEST, INBC.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001000; Glyco_hydro_10.
InterPro; IPR001547; Glyco_hydro_5.
Pfam; PF00942; CBM 3; 1.
Pfam; PF00341; Glyco_hydro_10.
Pfam; PF00313; Glyco_hydro 10; 1.
Pfam; PF00313; Glyco_hydro 10; 1.
PRINTS; PR00134; GLHYDRLASE10.
PRODOM; PD001947; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL HYDROL_F10; 1.
PROSITE; PS00659; GLYCOSYL HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Repeat; Multifunctional enzyme; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROTON DONOR (POTENTIAL).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; 0E0378171594DDAE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               479 QIGASNVTFKFVKLSSSVSGADYYLE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=92276330; PubMed=1592807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          285 285
792 792
1039 AA; 117641 M
                                                               EMBL; X13602; CAA31936.1; .-. PIR; S02711; S02711.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M76588; AAA22302.1; -. PIR; A41897; A41897.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1039
416
570
618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
nes 36; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=PL236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YCEA PAELA
AC PAERA
AC DAR PO
DT 16-APR
DT 16-APR
DT 16-APR
DE HYPOCH
DE PACHEL
CO BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTET
CON BACTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce) or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                    7 VQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAAMGC 66
                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               endo-beta-1,4-glucanase.";
J. Bacteriol. 174:3522-3531(1992).
-! CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
-! CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucans.
-! PTW: A SHORT FORM (EGA-S) ARISES FROM POSTTRANSLATIONAL
--- PTW: A SHORT FORM (EGA-S) ARISES FROM POSTTRANSLATIONAL
--- SIMILARITY: BELONGS TO CELLULASE FAMILY J (FAMILY 44 OF GLYCOSYL
--- HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hansen C.K., Diderichsen B., Joergensen P.L.; "celA from Bacillus lautus PL236 encodes a novel cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Paenibacillus lautus (Bacillus lautus).
Bacteria, Firmicutes, Bacillales, Paenibacillaceae, Paenibacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 41, Last annotation update)
Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase):
(Cellulase A) (EG-A).
                                                                                                                                                                                                                                                                                                                                                         3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BY SIMILARITY).
                                                                                                                                                                                                                                                                                        Length 145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or se...

EMBL, M76589, AAA...

R PIR, B41897, B41897.

DR HSSP, Q06851; INBC.

DR InterPro; IPR001956; CBD 3.

DR Pfam; PP00942; CBM 3; 1.

DR ProDom; PD001947; CBD 3; 1.

Cellulose degradation; Hydrolase; Glycosidase; Signal.

FT SIGNAL 34 700 CELLULOSE-BIDDING (BY SIMI. THAIN 34 700 CELLULOSE-BIDDING (BY SIMI. THAIN 31 313 BY SIMILARITY.

THAIN 313 BY SIMILARITY.

713 213 BY SIMILARITY.

76910 MW; 3D5C8CADA53EBE0F CRC64
                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                             SEQUENCE 145 AA; 15782 MW; 9514E3A71B106AEB CRC64;
                                                                                                                                                                                                                                                                                        Query Match 38.5%; Score 180.5; DB 1; Best Local Similarity 43.9%; Pred. No. 3.4e-13; Matches 36; Conservative 18; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      700 AA.
HSSP, Q06851; INBC.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; Cellul_bind.
Pfam; PF00942; CBM_3; 1.
ProDom; PD001947; CBD_3; 1.
Hypothetical protein.
NON TER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0 SNIQISFG--NHTGTNSDTYVE 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  67 GNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=92276330; PubMed=1592807;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=1401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=PL236;
```

```
RESULT 10
CIPB CLOTM
                                                                                                                  ద
                                                                                                                                                                               셤
                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                          1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                                                                          01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
(Cellulase I).
                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Gene sequence and properties of Cell, a family E endoglucanase from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Clostridium thermocellum.";
J. Gen. Microbiol. 138:307-316(1933).
-!- FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-GLUCANS. PRINCIPALLY ACTIVE AGAINST BARLEY BETA-GLUCAN.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucosis.
-!- PATHWAY: Cellulose degradation.
-!- SIMILARITY: Belongs to cellulase family E (family 9 of glycosylhydrolases).
                                                                                                                                                                                                                                                                                                                                                                  Bacteria; Firmicutes, Clostridia, Clostridiales, Clostridiaceae,
Clostridium.
                            4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY) .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hazlewood G.P., Davidson K., Laurie J.I., Huskisson N.S., Gilbert H.J.;
Pred. No. 2e-11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Prodom; PD001947; CBD 3; 1.7. PROSITE; PS00592; GLYCOSYL HYDROL F9 1; 1. PROSITE; PS00698; GLYCOSYL HYDROL F9 2; 1. Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35A60069A514A927 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULOSE-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ENDOGLUCANASE I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
       44.9%; Preu. ....
                                                                                                                                                                                                                               879 AA
                                                                                                                                       60 DWAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., AND SEQUENCE OF 56-69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSSP; P26221; 1TF4.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CelTul bind.
InterPro; IPR001701; Glyco_hydro_9.
InterPro; IPR008928; Glyco_trans_6hp.
Pfam; PF00942; CBM 3; 2.
Pfam; PF00759; Glyco_hydro_9; 1.
                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=93171873; PubMed=8436949;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    495 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; L04735; AAA20892.1; -. PIR; A47704; A47704.
                           40; Conservative
                                                                                                                                                                                                                                                                                                                                                     Clostridium thermocellum.
                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       448
486
         Best Local Similarity
Matches 40; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  879 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN=NCIB 10682;
                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=1515;
                                                                                                                                                                                                                             GUNI CLOTM
Q02934;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
                                                                                                                                                                                             엄
                                                                                                                  ઠે
                                                                                                                                                 셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                     36 IKGEVULQYANGNAGATSNSINPRFKIINNGTKAINLSDVKIRYYYTKEGGASQNFW-CD 794
                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tildentification of the cellulose-binding domain of the cellulosome subunit S1 from Clostridium thermocellum YS.";

ERMS Microbiol. Lett. 78:181-186(1992).

-!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE CELLULOLICSE TO THE CATALYTIC DOMAINS OF THE CELLULOSOME.

CELLULOLYTIC ENZYWES PROBABLY THROUGH THE BINDING OF THE NINE REPEATED DOMAINS WITH THE DUPLICATED SEGMENT OF 24 AMINO ACIDS PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSOME.

-!- SUBCELLULAR LOCATION: Cell surface.

-!- DOMAIN: THE COMPONENTS OF THE CELLULOSOME.

-!- DOMAIN: THE COMPONENTS OF THE CELLULOSOME.

-!- SIMILARITY: Contains at least 3 cohesin domains.
                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                          01-077-1994 (Rel. 30, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/SL)
(Cellulose integrating protein B) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clostridium thermocellum.
Bacteria, Firmicutes, Clostridia, Clostridiales, Clostridiaceae,
                                       1;
  Length 879;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=93146373; PubMed=1490597;
Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
Gilbert H.J.;
                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ProDom; PD001947; CBD 3; 1.

PROSITE; PS00018; EF HAND; UNKNOWN 1.

PROSITE; PS00448; CLŌS CELLULOSOME RPT; 2.

Cellulose degradation; Cell wall; Glycoprotein; Repeat.

NON TER <1 80 COHESIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COHESIN 1.
LINKER (PRO/THR-RICH)
  DB 1;
Ouery Match 35.3%; Score 165.5; DB 1
Best Local Similarity 33.0%; Pred. No. 1.4e-10;
Matches 29; Conservative 26; Mismatches 32
                                                                                                                                                                                                                                                                                                    772 AA
                                                                                                                                                                                    88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      entities requires a license agreement (Se
or send an email to license@isb-sib.ch).
                                                                                                                                                               61 WAAMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001956; CBD 3.
InterPro; IPR001956; Cellul bind.
InterPro; IPR002102; Cohesin.
InterPro; IPR002105; Dockerin.
InterPro; IPR002048; EF-hand.
InterPro; IPR002048; EF-hand.
Pfam; PF00964; CBM 3; 1.
Pfam; PF00964; Cohesin; 3.
Pfam; PF00404; Dockerin 1; 2.
                                                                                                                                                                                                                                                                                                                                           01-OCT-1994 (Rel. 30, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X68233; CAA48312.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Clostridium.
                                                                                                                                                                                                                                                                                                      CLOTM
                                                                                                                                                                                                                                                                                                                       201866;
```

```
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GRN
 277 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 335
                                                                                                                                    9
                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1820-1853 FROM N.A.
MEDLINE=33209931; PubMed=8458832;
Fujino T., Beguin P., Aubert J.-P.;
Fujino T., Beguin P., Aubert J.-P.;
Grganization of a Cloetridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface.";
[3]
[3]
                                                                                                              29; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                       "Sequencing of a Clostridium thermocellum gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97076134; PubMed=8918451;
Tormo J., Lamed R., Chixino A.J., Morag E., Bayer E.A., Shoham Y.,
                                                                                                                                                                                                                                                                                                                                         Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
NCBI_TaxID=1515;
                                                                                                                                                                                                                                                              01-JUN-1994 (Rel. 29, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Cellulosomal scaffolding protein A precursor (Cellulosomal glycoprotein S1/SL) (Cellulose integrating protein A) (Cohesin) CIPA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose."; EMBO J. 15:5739-5751(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                  Gerngross U.T., Romaniec M.P.M., Kobayashi T., Huskisson N.S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "A cohesin domain from Clostridium thermocellum: the crystal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              structure provides new insights into cellulosome assembly.";
Structure 5:381-390(1997).
                                                                                          Length 772;
                                                                      BBF06DE5E094FE10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43.
STRAIN=ATCC 27405 / DSM 1237;
MEDLINE=93302508; PubMed=8316083;
COHESIN 2.
LINKER (PRO/THR-RICH).
                   CELLULOSE-BINDING.
LINKER (PRO/THR-RICH).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361.
MEDLINE=98022914; PubMed=9402065;
                                                                                                                                                                                                336 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
                                                                                                                                                                           61 WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
                                                                                         35.2%; Score 165; DB 1; 36.1%; Pred. No. 1.4e-10; ive 23; Mismatches 29;
                                                                                                                                                                                                                                                   PRT; 1853 AA
                                       COHESIN 3.
DOCKERIN 1.
DOCKERIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=97238934; PubMed=9083107;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mol. Microbiol. 8:325-334(1993).
                                                                      82491 MW;
                                                                                                      36.1%;
                                                                                                                35; Conservative
                                                                                                                                                                                                                                                                                                                                   Clostridium thermocellum.
                                                                                                                                                                                                                                                   STANDARD;
240
272
4439
461
607
733
                                                                       772 AA;
                                                                                                     Local Similarity
94
241
273
440
710
743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Steitz T.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                               Demain A.L.;
                                                                                                                                                                                                                                                  CIPA CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                             homology.
                                                                                          Query Match
                                                                      SEQUENCE
DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                        DOMAIN
                                                                                                                                                                                                                                                            Q068<u>5</u>1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Frolow
                                                            DOMAIN
                                                                                                                Matches
SETTTTTS
                                                                                                                                                      셤
                                                                                                                                                                                                                                                            ð
                                                                                                                                                                           ò
                                                                                                                                                                                               셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
Tavares G.A., Beguin P., Alzari P.M.;
"The crystal structure of a type I cohesin domain at 1.7-A resolution.";
J. Mol. Biol. 273:701-713(1997)

- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE CELLULOLYTIC ENZYMES.

- SUBCELLULAR LOCATION: Cell surface.
- SUBCELLULAR LOCATION: Cell surface.
- DOMAIN: THE COHESIN DOMAINS BIND TO THE DOCKERIN DOMAIN BORN BY THE CATALYTIC COMPONENTS OF THE CELLULOSOME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    R InterPro; IPR002102; Cohesin.
R InterPro; IPR002102; Cohesin.
R InterPro; IPR002102; Cohesin.
R InterPro; IPR002102; Cohesin.
R Pfam; PF00942; Cohesin; 9.
R Pfam; PF00942; Cohesin; 9.
R Pfam; PF00944; Dockerin.; 9.
R ProDom; P0001194; EF HAND; UNKNOWN 1.
R PROSITE; PS00448; CLOS CELLULOSOME RPT; 2.
R PROSITE; PS00448; CLOS CELLULOSOME RPT; 2.
R PROSITE; PS00448; CLOS CELLULOSOME RPT; 2.
R Cellulose degradation; Cell wall; Glycoprotein; Repeat; Signal; 3D-structure.
I SIGNAL 1 28 CELLULOSOMAL SCAFFOLDING PROTEIN A. CHAIN 29 1853 COHESIN 1.
I DOMAIN 183 322 COHESIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER (PRO/THR-RICH).
CELLULOSE-BINDING (BY SIMILARITY).
LINKER (PRO/THR-RICH).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A -> AA (IN REF. 1)
                                                                                                                                                                                                                               -!- SIMILARITY: Contains 9 cohesin domains.
-!- SIMILARITY: Contains 2 dockerin domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COHESIN 5.
COHESIN 7.
COHESIN 8.
COHESIN 9.
DOCKERIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOCKERIN 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COHESIN 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, L08665; -; NOT ANNOTATED_CDS.
EMBL, X67506; CAA47840.1; -.
PIR, 336859; 336859.
PDB, 1ANU, 23-JUL-97.
PDB, 1ANU, 26-SEP-97.
INC, 26-SEP-97.
INC, 26-SEP-97.
INC, 26-SEP-97.
InterPro; IPR001956; Cellul_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COHESIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28
1853
182
322
363
522
525
559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1548
1791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1824
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59
                                                                                                                                                 STRAIN-ALTOSEPTICA FCBR C18;
MEDLINE-98299944; PubMed-9636315;
Olsen O., Thomsen K.K., Weber J., Duus J.O., Svendsen I., Wegener C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                     linkages in cellulose, lichenin and cereal beta-D-glucans.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                                                                von Wettstein D.;
"Transplanting two unique beta-glucanase catalytic activities into one multienzyme, which forms glucose.";
Biotechnology 14:71-76(1996).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V)
(Cellulase V).
                                5, Last sequence update)
1, Last annotation update)
cursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase N)
                                                                                   Erwinia carotovora.
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31.9%; Score 149.5; DB 1; Length 444; 38.6%; Pred. No. 3.8e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 444 ENDOGLUCANASE N.
168 PROTON DONOR (BY SIMILARITY)
255 UNCLEOPHILE (BY SIMILARITY).
48300 MW; FA7E4179004CBB43 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30; Indels
                                                                                                                                                                                                                                                                                                                                                                                        HSSP; 085465; LADA.
InterPro; IRR001956; CBD 3.
InterPro; IRR001956; CBD 3.
InterPro; IRR001956; CBD 3.
InterPro; IRR001547; Glyco_hydro_5.
Pfam; PP00150; cellulase; 1.
ProDom; P001547; CBD 3; 1.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Cellulose degradation; Hydrolase; Glycosidase; Signal.
31
  444 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   505 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            414 DWANVGPNNIVTSTGTPAASTDKANRYV 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 DWAAMGCGNIRASFGSVNPATPTADTYL
                      01-NOV-1997 (Rel. 35, Created)
                                                                                                                                                                                                                                                                                                                                                                              EMBL; L39788; AAC37033.1; -.
                                           28-FEB-2003 (Rel. 41, Last
Endoglucanase N precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                444 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                        hydrolases).
                                                                                                                    NCBI_TaxID=554;
                                                                 Cellulase N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Simil
Matches 34; C
                                 11-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GUNV ERWCA
Q47096;
  GUNN ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
ACT_SITE
ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
GUNV_ERWCA
                                                                            CELN.
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8844448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   365 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             35.2%; Score 165; DB 1; Length 1853; larity 36.1%; Pred. No. 3.9e-10; Conservative 23; Mismatches 29; Indels 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA; 196831 MW; 3ABDDC03ABFC5372 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||: ||: || ||: 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 WAAM-----GCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                       440
452
454
                                                                                                                                                                                                                                                                                                                                              319
3382
3382
400
403
                                                                                                                                                                                                                                     482
484
489
                                                                                                                                                                                                                                                                     Local Similarity
es 35; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1283
1289
1293
                                                                                                                                                                                 441
453
455
467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1313
1324
1329
1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1351
                                                                                                                                                                                                                                               483
                                                                                                                                                                                                                                                                                                                                                                                                  1249
1251
1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1306
                                                                                                                                                                                                                                                                      491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match
Best Local Si
Matches 35,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
SEQUENCE
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
STEAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
```

RESULT 12 GUNN_ERWCA

```
PIR; S54744; S54744.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     353
168
256
204 AA;
                                                                      SEQUENCE FROM N.A.
                                               NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT SITE
ACT SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GUX2 CL(
P50900;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLOSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUX2_
 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Endoglucanase V1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V1)
(Cellulase V1).
                                                                                                                                                     MOI. Gen. Genet. 241:341-350(1993).
-!- FUNCTION: Endoglucanase with some exoglucanase activity.
-!- FUNCTION: Endoglucanase with some exoglucanase activity.
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
!INKages in CATIVITY: Secreted.
-!- SUBCELLULAR LOCATION: Secreted.
-!- MISCELLANBOUS: Has a pH optimum of about 7.0 and a temperature optimum about 42 degrees Celsius.
-!- SIMILARITY: Belongs to cellulase family A (family 5 of glycosyl
                                                                                                                                carotovora: evidence for an evolutionary 'mix-and-match' of enzyme
                        Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LINKER.
CELLUJOSE BINDING (BY SIMILARITY).
RELOUDONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31.9%; Score 149.5; DB 1; Length 505;
                                                                                                         Cooper V.J.C., Salmond G.P.C.; "Molecular analysis of the major cellulase (CelV) of Erwinia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pram; Fruelso, Control of the Property of the Property of the Prostrat, Property of Cartons of the Prostrat of the Prostrat of the Prostration, Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54900 MW; DBEA9337BB4D2623 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Pred. No. 4.4e-09; 19; Mismatches 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ENDOGLUCANASE V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | | | | | | | | | : | | : | | : | | : | DWANVGPINIVTSTGTPAASTDKANRYV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                       EMBL, X76000; CAA53592.1; -.
PIR; 539962; 539962.
HSSP, 085465; 1A3H.
InterPro; IPR001956; CBD 3.
InterPro; IPR00156; CelTul bind.
InterPro; IPR001547; Glyco_hydro_5.
Pfam; PP00942; CBM 3; 1.
Pfam; PP00150; celTulase; 1.
                                 Enterobacteriaceae; Pectobacterium
                                                                                           MEDLINE=94067016; PubMed=8246888;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 38.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            505 AA;
            Erwinia carotovora
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                      hydrolases).
                                                NCBI_TaxID=554;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GUNW ERWCA
Q59395;
                                                                                                                                              domains.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQÜENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GUNW_ERWCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BE BE BE BE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isp-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDG--GSSTLVYNC
                                                                                                                                                                                                                                           Mae A., Heikinheimo R., Palva B.T., "Structure and regulation of the Erwinia carotovora subspecies carotovora SCC3193 cellulase gene celV1 and the role of cellulase in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5; Gaps
                                                                                                                                                                                                                                                                                                                                                          phytopathogenicity.";
Mol. Gen. Genet. 247:17-26(1995).
Mol. Gen. Genet. 247:17-26(1995).
-!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
-!--SuderEnLudian Locations, lichenin and cereal beta-D-glucans.
-!--SuderEnLudan Location: Secreted.
-!--SuderEnLudan Locations to cellulase family A (family 5 of glycosylhydrolases).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Exoglucanase II precursor (EC 3.2.1.91) (Exocellobiohydrolase II)
(1,4-beta-cellobiohydrolase II) (Avicelase II).
Bacteria, Proteobacteria, Gammaproteobacteria, Enterobacteriales;
Enterobacteriaceae, Pectobacterium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Clostridium stercorarium.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULOSE-BINDING (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        31.4%; Score 147.5; DB 1; Length 504; 37.5%; Pred. No. 7.4e-09; live 19; Mismatches 31; Indels 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTON DONOR (BY SIMILARITY).
NUCLEOPHILE (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0D7ECF74781565FA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSP, O85465, 1A3H.

InterPro; IRR001956; CBD 3.

InterPro; IRR001545; Cellul bind.

InterPro; IRR001547; Glyco_hydro_5.

Pfam; PF00942; CBM 3; 1.

Propon; PF00150; cellulase; 1.

Propon; PF001997; CBD 3; 1.

ROSITE; PS00659; GLYCOSYL HYDROL F5; 1.

Cellulose degradation; Hydrolase; Glycosidase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ENDOGLUCANASE V1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           914 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DWANVGPNNIVTSTGTPAASTDKANRYV 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 DWAAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CATALYTIC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LINKER
                                                                                                                                                                                STRAIN=SCC3193;
MEDLINE=95231512; PubMed=7715600;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54963 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X79241; CAA55823.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 37.5
les 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
```

```
OC NOBLITAZID=1510,
RM (TEST TAZID=1510),
RM (TEST TAZID=1510),
RM (TEST TAZID=1510),
RM (TEST TAZID=1510),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TEST TAZID=11754),
RM (TE
```

\$ B &

61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
||::| |: :| ||||
|822 WASIGSSNVTGTFVKMDGATTGADYYLE 849

Search completed: May 11, 2004, 12:07:29 Job time : 2.39913 secs

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q82QF2;
01-JUN-2003
01-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q82QF2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HODER REPAREMENT OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPER
                                                                                                                                                                                  May 11, 2004, 11:45:55; Search time 10.2683 Seconds (without alignments) 2704.020 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                              469
1 VSGCVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1017041
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1017041 segs, 315518202 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            summaries
                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_phage: * sp_plant: *
                                                                                                                                                                                                                                                                                                              US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                         BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                             Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                                                      Run on:
```

		Description	Q82qf2 streptomyce	Q7x2n2 thermomonos	Q9rfx6 caldibacill		Q91818 caldibacill		Q9z4il bacillus sp	Q9aqg4 caldicelluī	O24820 thermophili	Q9x3p5 caldicellul	Q9agg7 caldicellul	Q9x3p6 caldicellul	P96311 anaerocellu	Q9aqh0 caldicellul	052374 caldicellul	Q93tj6 bacillus su
SUMMARIES		QI .	Q82QF2	Q7X2N2	Q9RFX6	Q9RFX5	Q91818	Q8KKF7	092411	Q9AQG4	024820	Q9X3P5	Q9AQG7	Q9X3P6	P96311	ОЭАОНО	052374	093ТЛ6
		B	16	0	~	8	7	7	7	0	7	7	0	7	7	0	7	7
		Match Length DB	741	616	170	930	921	1091	997	1751	1000	1770	261	1426	1711	966	1779	499
	ouerv	Match	54.8	48.7	48.1	46.2	45.7	44.7	43.6	41.5	41.0	41.0	40.8	40.8	40.7	40.6	40.6	40.5
		Score	257	228.5	225.5	216.5	214.5	209.5	204.5	194.5	192.5	192.5	191.5	191.5	191	190.5	190.5	190
	Result	NO.	н	73	ო	4	2	9	7	80	6	10	11	12	13	14	15	16

- ααααααας υπα_υυυυας,,,	Q8/g89 girglymosto Q8/g82 bomo sapien Q9/32 homo sapien Q9/32 homo sapien Q8:zf2 homo sapien P9:1006 caenorhabdi Q8yru7 anabaena sp Q9gyg5 caenorhabdi Q99yg5 caenorhabdi
2 093LD0 2 045332 2 045332 2 045430 2 08RPQ6 2 09RPPQ6 2 09RPPQ6 2 09RPPQ6 2 09RPQ6 2 09RPQ9 2 09734 2 09734 2 04539 2 04539 2 046392 2 046392 2 093LI9 3 099LI9	13 Q8JGA9 4 Q9Y3Z2 2 Q9L3J2 4 Q94858 4 Q8IZF2 5 P91006 5 Q9GYGS 12 Q993M3
508 4499 5019 1919 4403 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646 11646	121 618 660 986 1300 1428 511 499
4 K W W W W W W W W W W W W W W W W W W	16.2 16.1 16.1 16.1 15.7 15.7 15.7
1990 186 186 186 172 172 172 155 157 150 153 130 130 117 119 119 117 119 119 119 119 119	76 75.5 75.5 75.5 75.5 73.5 73.5
1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	0 E E E E E E E E E E E E E E E E E E E

ALIGNMENTS

Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.;
Sakaki Y., Hattori M., Omura S.;
"Complete genome sequence and comparative analysis of the industrial microorganism Erreptomyces avermitilis.",
Mat. Biotechnol. 21:526-531(2003).
Mat. Biotechnol. 21:526-531(2003).
BMBL; ABOOSO32; BAC68267.1;
GO; GO:0005975; P:nydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001956; CBI 3.
InterPro; IPR001956; CBI 3.
InterPro; IPR001961; FN_III. SEQUENCE FROM N.A.
STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=21477403; PubMed=11572948;
Omura S., Ikeda H., Ishlakawa J., Hanamoto A., Takahashi C.,
Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
"Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary Streptomyces avermitilis. Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces. NCBI_TaxID=33903; STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165; MEDLINE=22608306; PubMed=12692562; 01-JUN-2003 (TrEMBLrel. 24, Created) 01-JUN-2003 (TrEMBLrel. 24, Last sequence update) 01-OCT-2003 (TrEMBLrel. 25, Last annotation update) Putative cellulose 1,4-beta-cellobiosidase. GUXAI OR SAV557. metabolites."; Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001). 741 AA PRT; SEQUENCE FROM N.A.

```
20 GSLVVQYRAADTNAGDNQLKPHFRIVNRGTSSVPLSELTIRYWYTVD-GDKPQVFNCDWA 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Caldibacillus.
NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Gaps
                                         MEDLINE=20120520; PubMed=10653733; Melson P.J., Bergquist.P.L.; Should A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist.P.L.; Sunan A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist.P.L.; Agne encoding a novel multidomain beta-1,4-mannanase from kraft pulp., Microbiol. 66:664-670(2000). EMBL; AF163837; AAF22273.1; -... HSSP; Q0681; NRC.

HSSP; Q0681; NRC.

InterPro; IPR001956; CBD3.

InterPro; IPR001956; CBD3.

InterPro; IPR001956; Cellul_bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE=20120520; PubMed=10653733;
Sunna A., Gibbs M.D., Chin C.W.J., Nelson P.J., Bergquist P.L.;
sunna a., Gibbs a novel multidomain beta-1,4-mannanase from
Caldibacillus cellulovorans and action of the recombinant enzyme on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 930;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33 POTENTIAL.
930 MULTIDOMAIN BETA-1,4-MANNANASE.
101576 MW; 0086638D54D1A2CC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
48.1%; Score 225.5; DB 2; Length
Best Local Similarity 47.7%; Pred. No. 8.2e-17;
Matches 41; Conservative 17; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00942; CBM 3; 1. Probom; PD001947; CBD 3; 1. Hypothetical protein.
NON TER 1 1 SEQÜENCE 170 AA; 18493 MW; 7AC9D33F44E3A0B4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY_2000 (TrEMBLrel. 13, Created)
1-MAY_2000 (TrEMBLrel. 13, Last sequence update)
01-0CT-2003 (TrEMBLrel. 25, Last annotation update)
Multidomain beta-1,4-mannanse precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               46.2%; Score 216.5; DB 2; 46.5%; Pred. No. 6.2e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Appl. Environ. Microbiol. 66:664-670(2000).
EMBL, AF163837, AAF22274.1; -.
HSSP, Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProDom; PD001947; CBD_3; 2.
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001956; CBD 3.
InterPro; IPR008965; Cellul bind.
InterPro; IPR004302; Chitin binding 3.
InterPro; IPR001547; Glyco hydro 5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 AMGCGNIRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00150; cellulase; 1.
Pfam; PF03067; Chitin_bind_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00942; CBM 3; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                930 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9RFX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O9RFX5
                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DDT TALE REAL SOLUTION OF THE STANK AND THE SOLUTION OF THE STANK AND THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLUTION OF THE SOLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Hypothetical protein (Fragment).
Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Caldibacillus.
NCBI_TAXID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=TMS1;
Posta K., Beki E., Kukolya J., Hornok L.;
Posta K., Beki E., Kukolya J., Hornok L.;
"Phylogenetic relationships of Tf cel5B, a new endoglucanase encoding gene from Thermobifida fusca.";
Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
EMBL; AY298814; AAP56348.1; -.
SEQUENCE 616 AA; 67701 MW; 24FFCIEAIA3F5639 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 48.7%; Score 228.5; DB 2; Length 616; Best Local Similarity 52.3%; Pred. No. 1.8e-16; Matches 45; Conservative 16; Mismatches 22; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                 54.8%; Score 257; DB 16; Length 741; 56.3%; Pred. No. 1.7e-19; ive 13; Mismatches 25; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                                                                        741 AA; 77396 MW; 79404B40B2B4A7AF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Streptosporangineae, Nocardiopsaceae, Thermobifida
NCBI_TaxID=2021;
InterPro; IPR008957; FN III-like.
InterPro; IPR001524; Glyco_hydro_6.
Pfam; PP00942; CBM 3; 1.
Pfam; PP00134; Glyco_hydro_6; 1.
Pfam; PP01341; Glyco_hydro_6; 1.
ProDom; PD001947; CBD 3; 1.
ProDom; PD001947; CBD 3; 1.
ProDom; PD001947; CBD 3; 1.
ProDom; PD0055; Glyco_hydro_6; 1.
Complete Proceome:
SMART; SM00060; FN3; 1.
Complete Proceome:
SEQUENCE 741 AA; 77396 MW; 79404B40B2B4AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  616 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 170 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          651 AAVGCSNVSLKVVPLTTPVPGADAYLE 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62 AAMGCGNIRASFGSVNPATPTADTYL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2003 (TrEMBLrel. 25, 01-OCT-2003 (TrEMBLrel. 25, 01-OCT-2003 (TrEMBLrel. 25,
                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 56.3%
nes 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Thermomonospora fusca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Endoglucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9RFX6
```

07X2N2

OGREXE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREY
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREYE
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
OGREY
O

ઠ ద 8

datches

8 g 8 셤

8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8 \pm 8

Matches

ò

```
Pastor F.I.J., Pujol X., Blanco A., Vidal T., Torres A.L., Diaz P., "Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its performance in pulp refining.";
                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                    STRAIN=BP-23;
Sanchez M.M., Pastor F.I.J., Diaz P.;
"Paenibacillus sp. BP-23 family 48 cellulase. Cloning and performance
                                                  Paenibacillus sp. BP-23.
Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appl. Microbiol. Biotechnol. 55:61-68(2001).
-!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
EMBL; AJ113164; CAB38941.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match

44.7%; Score 209.5; DB 2; Length
Best Local Similarity 46.0%; Pred. No. 4.4e-14;
Matches 40; Conservative 21; Mismatches 25; Indels
Cellulose 1,4-beta-cellobiosidase precursor (EC 3.2.1.91)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bacillus sp. BP-23.
Bacteria, Firmicutes, Bacillales, Bacillaceae, Bacillus.
NCBI_TaxID=89769;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GO; GO:0008810; F:cellulase activity; IEA. GO; GO:0005215; F:transporter activity; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-1999 (TrEMBLrel. 10, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=BP-23;
MEDLINE=21129642; PubMed=11234960;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-1999 (TrEMBLrel. 10, Last e 01-OCT-2003 (TrEMBLrel. 25, Last e Cellulase precursor (EC 3.2.1.4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                    NCBI_TaxID=198119;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9Z4I1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
  SO CREATE BEAR AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CREATE AND CRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                    3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 GGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWA
  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Caldibacillus cellulovorans.
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=20519260; PubMed=11065373;
Sunna A., Gibbs M.D., Bergquist P.L.;
"A novel thermostable multidomain 1,4-beta-xylanase from
'Caldibacillus cellulovorans' and effect of its xylan-binding domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Microbiology 146:2247-2955 (2000).
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HIGH AFZO03043, AAFG1649.1; -.

REMBI, AFZO030453, AAFG1649.1; -.

R GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . O:

R GO; GO:0005975; P:carbohydrate metabolism; IEA.

R InterPro; IPR001956; CBD 3.

R InterPro; IPR001956; CBIL1 bind.

R InterPro; IPR00190979; Gal bind like.

R InterPro; IPR001000; Glyco_hydro_lo.

R Pfam; PF00191; CBM 3; 2.

R Pfam; PF00131; Glyco_hydro 10; 1.

R Pfam; PF0031; Glyco_hydro 10; 1.

R Probom; PB001947; CBD 3; 2.

R Probom; PB001947; CBD 3; 2.

R Probom; PB001947; CBD 3; 2.

R Glycosidase; Hydrolase; Synal; Xylan degradation.

R SIGNAL

R SIGNAL
  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 45.7%; Score 214.5; DB 2; Length 921; Best Local Similarity 46.5%; Pred. No. 1e-14; Matches 40; Conservative 16; Mismatches 29; Indels 1;
  28; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNAL 1 33 POTENTIAL.
CHAIN 34 921 BETA-1,4-XYLANASE XYNA.
SEQUENCE 921 AA; 102380 MW; C5DDD1A7F7567413 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                    01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Beta-1, 4-xylanase XynA precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 22, Created)
(TrEMBLrel. 22, Last sequence update)
(TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                        921 AA
17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  830 WVGCSNLRGSLVKLTTGRTGADYYLE 855
                                                                                                                                                                                            :|| |:| || ::
839 QVGCSNLRGSFVKLSTGRTGADYYIE 864
                                                                                                                                                              63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63 AMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                        PRT;
40; Conservative
                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           on enzyme activity.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=74586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-OCT-2002
01-OCT-2002
01-OCT-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QBKKF7
QBKKF7;
                                                                                                                                                                                                                                                                                                                                                  Q9L8L8
```

RESULT 6 Q8KKF7

ద

à

61

1; Gaps

Length 1091;

Gaps

```
678 GVKVLYKNNETSASTGSIRPWFKIVNGGSSSVDLSRVKIRYWYTVDGDKPQSAV-CDWAQ 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hydrolyzing O-glycosyl . . .; IEA.
                                                                                                                                                                                                   4 GVKVOYKNNDSAPGDNOIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Miyake K., Machida Y., Hattori K., Iijima S.; "Characterization of a multi-domain cellulase from an extremely thermophilic anaerobe strain NA10.", Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
                                             Length 1751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
41.0%; Score 192.5; DB 2; Length
Best Local Similarity 43.7%; Pred. No. 2.9e-12;
Matches 38; Conservative 17; Mismatches 31; Indels
                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1000 AA; 113265 MW; B9F659A56A752C6B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Beta-glucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GO; GO: 0003779; F: actin binding; IEA.
GO; GO: 0004553; F: hydrolase activity, hydrolyzing
GO; GO: 0004553; F: hydrolase activity, hydrolyzing
GO; GO: 000575; P: carbohydrate metabolism; IEA.
InterPro; IPR001589; Actbind actinin.
InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 3.
InterPro; IPR001000; Glyco hydro 10.
InterPro; IPR001547; Glyco hydro 10.
InterPro; IPR001547; Glyco hydro 10.
InterPro; IPR001547; Glyco hydro 10.
IPR00131; Glyco hydro 10; 1.
IPR00131; Glyco hydro 10; 1.
IPR00131; Glyco hydro 10; 1.
IPR00131; Glyco hydro 10; 1.
IPR051TE; PS00019; ACTININ 1; 1.
IPR051TE; PS00059; Glyco 10; 1.
IPR051TE; PS0059; Glyco IV; IPRDCI FF; 1.
IPR051TE; PS0059; Glyco IV; IPRDCI FF; 1.
                                             DB 2;
                                         Query Match
41.5%; Score 194.5; DB 2
Best Local Similarity 44.7%; Pred. No. 3.4e-12;
Matches 38; Conservative 14; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1000 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 1770 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | :| |: | |: | |: 431 AOIGASNVTFKFVKLSSSVSGADYYLE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 AAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                        : | : : | : : | 137 IGASIVITENFVKLSSGVSGADYYLE 761
                                                                                                                                                                                                                                                                                                                                                                  64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HYDROLASES).
EMBL; AB008029; BAA22939.1; -.
HSSP; Q06851; INBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              thermophilic anaerobe NA10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=67756;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=NA10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                024820
024820;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9X3P5
Q9X3P5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   024820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9X3P5
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESTRAINEL FKOTA N.A.

RESTRAINEL FKOTA N.A.

REDINEL STRONGE FKOTA N.A.

Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,

Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,

Reighs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,

RA Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,

RA GIBBS M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,

RY Thermophile Caldicellulosiruptor isolate Tok7B.1.";

REMEL, AROT8042; AAK06394.1; -

REMEL, AROT8042; ARAK06394.1; -

ROS GO.0005315; Fitransporter activity, hydrolyzing O-glycosyl ..; IEA.

GO; GO.0005515; Fitransporter activity, IEA.

GO; GO.0005315; Fitransport iEA.

ROS GO.0006810; Pitransport; IEA.

ROS GO.0006810; Pitransport; IEA.

RICEPTO; IPR001956; CelTul bind.

InterPro; IPR001956; CelTul bind.

InterPro; IPR001956; CelTul bind.

InterPro; IPR001956; CelTul cytFABP.

RAM: PR00179; Glyco hydro 9; 1.

PROSITE; PR00698; GLYCOSYL HYDROL F9 1; 1.

RROSITE; PR00698; GLYCOSYL HYDROL F9 2; 1.

RROSITE; PR00115; LIPCOSYL HYDROL F9 2; 1.

RROSITE; PR00115; LIPCOSYL HYDROL F9 2; 1.

RROSITE; PR00115; LIPCOSYL HYDROL F9 2; 1.

RROSITE; PR00115; LIPCOSYL HYDROL F9 2; 1.

RROSITE; PR00115; LIPCOSYL HYDROL F9 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CelB.
Caldicellulosiruptor sp. Tok7B.1.
Bacteria, Firmicutes, Clostridia, Clostridiales, Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                43.6%; Score 204.5; DB 2; Length 997;
43.7%; Pred. No. 1.4e-13;
Live 21; Mismatches 27; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE 997 AA; 106927 MW; F20CB119D6410629 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
    GO:0005975; P:carbohydrate metabolism; IEA.
GO:0006810; P:transport; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1751 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PSO0592; GLYCOSYL HYDROL F9 1; 1. PROSITE; PSO0698; GLYCOSYL HYDROL F9 2; 1. PROSITE; PS00213; LIPOCALIN; 1. Glycosidase; Hydrolase; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                  InterPro; IPR001956; CBD 3.
InterPro; IPR001956; CBD 11.
InterPro; IPR001956; Celliul bind.
InterPro; IPR001956; Celliul bind.
InterPro; IPR001956; FW III-like.
InterPro; IPR001971; fJyco hydro 9.
InterPro; IPR001951; Glyco Lrans 6hp.
InterPro; IPR000566; Lipocln_cytFABP.
Pfam; PF00942; CBM 3; 2.
Pfam; PF00041; fin3; 2.
Pfam; PF001947; CBD 3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 43.7% tes 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SM00060; FN3; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=80339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9AQG4;
01-JUN-2001 (
01-JUN-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09A0G4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8

COADAGA

COADAGA

COADAGA

COADAGA

COADAGA

COADAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

COAGA

C
DAR WEEN TO BE A SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY SECONDARY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                        Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme
       Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Caldicellulosiruptor sp. Tok/B.1.
Bacteria, Firmicutes, Clostridia, Clostridiales, Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L., "Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           R GO; GO:0003779; F:actin binding; IEA.

R GO; GO:0003779; F:actin binding; IEA.

R GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; I

R GO; GO:0005975; P:carbohydrate metabolism; IEA.

R InterPro; IPR001589; Actbind_actnin.

R InterPro; IPR001956; CBD interPro; IPR001906; CBD interPro; IPR001000; Glyco_hydro_5.

R InterPro; IPR001000; Glyco_hydro_5.

R Pfam; PF00942; CBM 3; 3.

R Pfam; PF00131; Glyco_hydro_10; I.

R Pfam; PF00131; Glyco_hydro_10; I.

R PRONITS; PR00134; GLHYDRLASE10.

R SWART; SW00633; Glyco_10; 1.

R PROSITE; PS00019; ACTINN 1; 1.

R PROSITE; PS00019; ACTINN 1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Curr. Microbiol. 40:333-340(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 261;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                             261 261
261 AA; 28759 MW; 4771744A26A6AE04 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                      thermophile Caldicellulosiruptor isolate Tok7B.1.
Curr. Microbiol. 40.333-340(2000).
EMBL; AF078040; ARK06391.1; -.
HSSP; Q06851; INBC.
GO; GO:0016787; F:hydrolase activity; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001955; CBD 3.
InterPro; IPR001955; CBD 3.
ProDom; PD001947; CBM 3; 1.
ProDom; PD001947; CBD 3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
40.8%; Score 191.5; DB 2;
Best Local Similarity 43.5%; Pred. No. 7.3e-13;
Matches 37; Conservative 15; Mismatches 32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 1426 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           177 IGASNVTFNFVKLSSGVSGADYYLE 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AF078737; AAD30364.1; -. HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-OCT-2003 (TrEMBLrel. 25,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HYDROLASES) .
                                                                                                                                                                                                                                                                                                                                       NON TER
NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                    Hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09X3P6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09X3P6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
Q9X3P6
       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GO; GO:0030246; F:carbohydrate binding; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
GO; GO:0005975; F:carbohydrate metabolism; IEA.
InterPro; IPR001556; CBD 3.
InterPro; IPR005584; CBD 1V.
InterPro; IPR005884; CBM 6.
InterPro; IPR005894; CBM 6.
InterPro; IPR008965; Cellul bind.
InterPro; IPR008965; Cellul bind.
InterPro; IPR008965; Gallul bind.
InterPro; IPR008979; Gallul bind.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                            Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P., Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
-!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
61y-Coryl hydrolase 6 (Fragment).
Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae; Caldicellulosiruptor.
NCBL TaxID=80339;
                                                                                                                  Caldicellulosiruptor sp. Tok7B.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41.0%; Score 192.5; DB 2; Length 1770;
43.5%; Pred. No. 5.7e-12;
tive 16; Mismatches 31; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1770 AA; 193641 MW; 8BAF1937D4926C92 CRC64;
01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00942; CBM 3; 3. Pfam; PF02018; CBM 4 9; 2. Pfam; PF02018; CBM 4 9; 2. Pfam; PF00312; CBM 4 9; 2. Pfam; PF0031; GlyCo hydro 10; 1. PRNTS; PR00134; GLHYDRLASE10. PRODOM; PM01947; CBD 3; 3. SMART; SM00606; CBD IV, 1. SMART; SM00603; GlyCo 10; 1. PR0SITE; PS00591; GLYCOSYL, HYDROL_F10; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1163 IGASNVTFNFVKLSSGVSGADYYLE 1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                              STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [1] __
SEQUENCE FROM N.A.
STRAIN-TOA7B.1,
MEDLINE-20171169, PubMed=10706665,
                                                                                                                                                                                                                                                                                                                                                                                                                                                          HYDROLASES).
EMBL; AF078737; AAD30363.1; -.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 43.5
nes 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Glycosidase; Hydrolase
                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                            NCBI TaxID=80339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
    DORRAND DRANGE DORRAND DRANGE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE SERVICE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
```

•

63

Gaps

1;

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQÜENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON TER
                                                     Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
  RESULT 14
Q9AQH0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |:|| :|| | |: | :: | |: | |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: || |: |
                                                                                                                                                                                                                           1 VSGG-VKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNC 59
                                                                                                                                                                                                    GVKVQYKUNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; II GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; II GO; GO:0005975; P:carbohydrate metabolism; IEA.

R InterPro; IPR001956; CBD 3.

R InterPro; IPR001956; CBD 3.

R InterPro; IPR001956; Glyco_hydro_9.

R InterPro; IPR001928; Glyco_hydro_9.

R Pfam; PF00191; Glyco_hydro_9; 1.

R Pfam; PF00191; Glyco_hydro_9; 1.

R PRINTS; PR00844; GLHTDRLASE40.

R PRODOm; PD011903; GLYCOSYL_HYDROL_F9_1; 1.

R PROSITE; PS00592; GLYCOSYL_HYDROL_F9_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDILINE-99154434; PubMed-9493383;

Zverlov V., Mahr S., Riedel K., Bronnenmeier K.;

Zverlov V., Mahr S., Riedel K., Bronnenmeier R.;

"Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile Anaerocellum thermophilum with separate glycosyl hydrolase family 9 and 48 catalytic domains.";

Microbiology 144:457-465(1998).

PIR: 731337; 731337.
                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Pirmicutes; Clostridia; Clostridiales; Anaerocellum group; Anaerocellum.
                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5
                                                                                              40.8%; Score 191.5; DB 2; Length 1426; 43.5%; Pred. No. 5.7e-12; ative 15; Mismatches 32; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40.7%; Score 191; DB 2; Length 1711; 43.8%; Pred. No. 8e-12; ive 18; Mismatches 30; Indels
PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
Glycosidase; Hydrolase.
SEQUENCE 1426 AA; 157544 MW; 29B3FDBB5D09A863 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1711 AA; 189979 MW; E3E987CEB9CD0C21 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. 03, Last sequence update) (TrEMBLrel. 25, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1711 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 742 DWAQIGASNVTFKFVKLSSSVSGADYYLE 770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                            : | |: : | |: 472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                                                                                                                                                                                                                                   64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1997 (TrEMBLrel. 03, Created) 01-MAY-1997 (TrEMBLrel. 03, Last seq 01-OCT-2003 (TrEMBLrel. 25, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1,4-beta-glucanase (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       43.8%;
                                                                       Query Match
Best Local Similarity 43.5%,
-hes 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Anaerocellum thermophilum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              39; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; T31337; T3133
HSSP; P26221; 1TF4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=31899;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P96311;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P96311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
  SOR
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                ₽
                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
63
                  09AQH0;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Glycosyl hydrolase 5 (Fragment).
Caldicellulosiruptor sp. Tok/PB.1.
Bacteria; Firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
                                                                                                                                                                                                                                                                      STRAIN=TOK7B.1;
MEDLINE=20171169; PubMed=10706665;
Gibbs M.D., Reeves R.A., Farrington G.K., Anderson P., Williams D.P.,
Bergquist P.L.;
"Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile Caldicellulosiruptor isolate Tok7B.1.";
Curr. Microbiol. 40:333-340(2000).
EMBL; AF078038; AAK06388.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caldicellulosiruptor sp. Rt69B.1.
Bacteria, firmicutes; Clostridia; Clostridiales; Syntrophomonadaceae;
Caldicellulosiruptor.
NCBI_TaxID=70295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KT691.1.";
Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
S-I- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=Rt69B.1;
Morris D.D., Gibbs M.D., Ford M., Thomas J., Bergquist P.L.;
"Family 10 and 11 xylanase genes from Caldicellulosiruptor sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40.6%; Score 190.5; DB 2; Length 996; 43.5%; Pred. No. 4.8e-12; Live 15; Mismatches 32; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               996 AA; 108275 MW; 3C72B6ED22F3C614 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-1998 (TrEMBLrel. 06, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Family 10 xylanase (EC 3.2.1.8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP, Q06851, 1NBC.
GO; GO:0016787; F:hydrolase activity; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001956; CBD 3.
InterPro; IPR008965; Celīul_bind.
InterPro; IPR008966; GH_BNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1779 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :| |: :| |: 003 IGASNVTFNFVKLSSGVSGADYYLE 927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 MGCGNIRASFGSVNPATPTADTYLQ 88
  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HYDROLASES).
EMBL; AF036924; AAB95326.1;
PIR; T31085; T31085.
HSSP; Q06851; 1NBC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF02012; BNR; 9.
Pfam; PF00942; CBM 3; 1.
ProDom; PD001947; GBD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 43.5% tes 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            966
```

```
4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA 63
       F:carbohydrate binding; IEA.
F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
P:carbohydrate metabolism; IEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match

40.6%; Score 190.5; DB 2; Length 1779;
Best Local Similarity 43.5%; Pred. No. 9.6e-12;
Matches 37; Conservative 14; Mismatches 33; Indels 1;
GO; GO:0030246; F:carbobydrate binding; IEA.
GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl
GO; GO:0005975; P:carbobydrate metabolism; IEA.
InterPro; IPR001584; CBD_3.
R InterPro; IPR001584; CBD_6.
R InterPro; IPR0018979; CBM_6.
InterPro; IPR0018979; CBM_6.
R InterPro; IPR0018979; Galbind.
R InterPro; IPR0018979; Galbind.
R InterPro; IPR0018979; Galbind.
R Pfam; PF02018; CBM_4 9; 2.
R Pfam; PF0318; CBM_4 9; 3.
R Pfam; PF0318; CBM_6; 1.
R Pfam; PF0318; CBM_6; 1.
R Pfam; PF0318; CBM_6; 1.
R Pfam; PF0318; CBM_6; 1.
R PRINTS; PR00134; GLHYDRALSE10.
R PRODON; PR00134; GLHYDRALSE10.
R SMART; SM0066; CBD IV; 1.
R SMART; SM00613; GLHYCOSYL HYDROL F10; 1.
R Glycosidase; Hydrolase; Xylan degradation.
O SEQUENCE 1779 AA; 194304 MM; CES269B6806B5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : | |: : | |: 1172 IGASINUTENFUKLISGVSGADYYLE 1196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MGCGNIRASFGSVNPATPTADTYLQ 88
         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
```

Search completed: May 11, 2004, 12:09:29 Job time : 11.2683 secs

```
RESULT 1
US-09-136-574A-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 43, Appl
Sequence 2, Appli
Sequence 1, Appli
Sequence 9, Appli
Sequence 4, Appli
Sequence 10, Appli
Sequence 12, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 7, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Appli
Appli
Appli
Appli
Appli
                                                                                    May 11, 2004, 12:04:16 ; Search time 4.31843 Seconds (without alignments) 1052.023 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 22,
Sequence 26,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 10,
Sequence 12,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                  /cgn2_6/ptodata/2/jaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/jaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/2/jaa/PCTUS_COMB.pep:*
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PCT-US95-13813-9
US-09-339-1598-4
US-09-198-955-110
US-09-694-531-12
US-09-670-141-10
US-10-72-12-12
US-08-460-462-2
US-08-460-455-2
US-08-460-455-2
US-08-460-455-2
US-08-460-455-2
US-08-30-394A-2
US-08-9006-636-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-136-574A-44
US-09-136-574A-43
US-07-862-588B-2
US-09-033-537A-1
                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -09-609-161B-22
                                                                                                                                                                                                                                                       389414 segs, 51625971 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                         protein search, using sw model
                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                       Issued Patents AA:*
                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                 US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Match
                                                                                                                                               Title:
Perfect score:
                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score
                                                          protein
                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                         Database
                                                                                      Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
```

```
Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions
                                       30, Appl
18, Appl
18, Appl
4640, Appl
2, Appl
38, Appl
31, Appl
18, Appl
19, Appl
19, Appl
19, Appl
26, Appl
68, Appl
68, Appl
68, Appl
68, Appl
68, Appl
                                                                                                                     Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                    Sequence
Sequence
Sequence
                                                          Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                 Sequence
                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STREET: Spring House Corporate Center, P.O. CITY: Spring House CATATE: PA COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:

MEDIUM TYPER. DISKETCE

COMPUTER: 18M COMPATIBLE

COMPUTER: 18M COMPATIBLE

COMPUTER: 18M COMPATIBLE

COMPUTER: 18M COMPATIBLE

COMPUTER: 19M DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <UNKNOWN:

PRIOR APPLICATION OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIMENT OF THE SEPTIME
                                 US-09-301-593-30
US-09-341-597-3
US-09-301-593-18
US-09-301-593-43
US-09-301-593-43
US-09-110-517-2
US-09-423-449-38
US-09-687-538B-8
US-09-687-538B-8
US-09-687-538B-8
US-09-687-538B-8
US-09-687-538B-8
US-09-421-439-32
US-08-481-995-68
US-08-481-995-68
US-08-481-995-68
US-08-481-995-68
US-08-481-995-68
                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 47, Application US/09136574A
Patent No. 6294366
GENERAL INFORMATION:
APPLICANT: Farrington, Graham K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 616 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 4:
CORRESPONDENCE ADDRESS
```

```
Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
  578 GVKVLYKNNETSASTGSIRPWFKIVNGGSSSVDLSRVKIRYWYTVDGDKPQSAV-CDWAQ 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STREET: Spring House Corporate Center, P.O. Box 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40.8%; Score 191.5; DB 3
43.5%; Pred. No. 1.8e-12;
tive 15; Mismatches 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: «Unknown»
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 08/932,571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: linear

MOLECULE TYPE: No. 6294366e

SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-136-574A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472 IGASNVTFNFVKLSSGVSGADYYLE 496
                                                                                737 IGASNVTFNFVKLSSGVSGADYÝLE 761
                                          64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 MGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                 Graham K.
                                                                                                                                                                                 Sequence 43, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 1426 amino acids
                                                                                                                                                                                                                                                                                                                                                                     Williams, Diane P. TITLE OF INVENTION: Compositi
                                                                                                                                                                                                                                                               Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: Spring House
STATE: PA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 19477
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                              APPLICANT: Farrington,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 43.5%
These 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: <Unknown>
N FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR
                                                                                                                                                                JS-09-136-574A-43
                                                                                                                                             RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
  셤
                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
                                                                                                                        9
                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDWAA
                                                                                  Gaps
                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 41.5%; Score 194.5; DB 3; Length 1751; Best Local Similarity 44.7%; Pred. No. 1.1e-12; Matches 38; Conservative 14; Mismatches 32; Indels 11;
                                        Length 616;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: Spring House Corporate Center, P.O. Box 457
                                                                                34; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Williams, Diane P. TITLE OF INVENTION: Compositions and Methods for
                                        DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
                                      ; Score 195.5; DB 3
; Pred. No. 2.4e-13;
15; Mismatches 34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                Sequence 44, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Farrington, Graham
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 1751 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                              Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEPHONE: 215-540-9200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION FOR SEQ ID NO: 44:
                                      41.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 19477
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino acid
                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'ELEX: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                           Best Local Similarity
Matches 38; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF
                                                                                                                                                                                                                                                                                                        RESULT 2
US-09-136-574A-44
US-09-136-574A-47
                                        Query Match
                                                                                                                                                            a
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                          a
```

```
3: Browdy and Neimark
419 Seventh Street N.W., Ste. 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         461 AQIGASNVSAAF -- ANFTGSNTDTYVE 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4492.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 9, Application PC/TUS9513813 GENERAL INFORMATION:
: 405 Lexington Avenue
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET UNMBER: 44
TELECOMMUNICATION:
TELEPHONE: 212-867-013
                                                                                                                                               IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH: 551 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                  Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: 212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SS: single
linear
                                                                                   ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 415
                                            10174 U.S.A.
                                                                                                                                                 COMPUTER: IBM COOPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-033-537A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US95-13813-9
                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                        GENERAL INFORMATION:
APPLICANT: Sch lein, Martin
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59580830 No. 5958083disk of No. 5958083th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNND-SAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWPTRDGGSSTLVYNC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 36.7%; Score 172; DB 2; Length 700; Best Local Similarity 44.9%; Pred. No. 1.1e-10; Matches 40; Conservative 18; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19-JAN-1990
PRIOR APPLICATION NUMBER: DET/DK91/00013
APPLICATION NUMBER: 19-JAN-1991
ATPORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
REGISTRATION NUMBER: 30,335 / 33,728
REGISTRATION NUMBER: 30,335 / 33,728
TELERRENCE/DOCKET NUMBER: 3425.204-US
TELERRAK: 212 867 0123
TELERRAK: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/09033537A
Patent No. 5558083
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Fich, Merete
APPLICANT: Toft, Annette Hanne
APPLICANT: Sh lein, Martin
TITLE OF INVENTION: Prevention Of Back-Staining
TITLE OF INVENTION: In Stone Washing
                                                                                                                                                                                                                                    STREET: 405 Lexington Avenue, 62nd floor CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||| :| || :| || :| || :| || 608 DWAKLGGSNIQISFGNHNGA--DSDTYAE 634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                 Sequence 2, Application US/07862588B Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IELERX: 212 00. 100: 2 INFORMATION FOR SEQ ID NO: 2 SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                               COMPUTEY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , MOLECULE TYPE: protein US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                 STATE: New York COUNTRY: U.S.A.
US-07-862-588B-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-033-537A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 엄
```

```
2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 35.3%; Score 165.5; DB 2; Length 551; Best Local Similarity 37.9%; Pred. No. 4.1e-10; Matches 33; Conservative 22; Mismatches 29; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Ramot University Authority for Applied
APPLICANT: Research and Industrial Development Ltd.
APPLICANT: Technion Research and Development Foundation Ltd.
APPLICANT: Bayer, Edward A.
APPLICANT: Bayer, Edward A.
APPLICANT: Wilchek, Meir
APPLICANT: Wilchek, Meir
APPLICANT: Wilchek, Meir
APPLICANT: Lamed, Raphael
APPLICANT: Shoham, Yuwal
TITLE OF INVENTION: MODIFIED CELLULOSE-BINDING DOWAIN (CBD)
TITLE OF INVENTION: PROTEINS AND USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Yeda Research and Development Co. Ltd
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,537A
FILING DATE: 02-MAR-1998
CLASSIFICATION OATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: 0993/95
FILING DATE: 03-SEP-1996
ATTORNEX/AGBNT INFORMATION:
NAME: Green, Reza
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 VSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFW-CD 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 35.2%; Score 165; DB 5; Length 167; Best Local Similarity 36.1%; Pred. No. 1.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Score 165; DB 4; Length 47; Pred. No. 3.9e-10; 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 100
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13813
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Schulein, Martin
APPLICANT: Schnorr, Kirk
APPLICANT: Andersen, Lene
APPLICANT: Bjornvad, Mads
TILE OF INVENTION: No. 6566114el Marmanases
FILE REFERENCE: 5440.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION NUMBER: US/09/339,159B
CURRENT FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 4, Application US/09339159B
Patent No. 656614
GENERAL INFORMATION:
APPLICANT: Kauppinen, Markus
                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: BAYE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35.2%;
36.1%;
                                                                                                                                                                                                                                                                                                                                                                                            LENGTH: 167 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 35.2%
Best Local Similarity 36.1%
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT ORGANISM: Bacillus
                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US95-13813-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
US-09-339-159B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-339-159B-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 <sup>o</sup>D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

RESULT

```
1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10; Gaps
                                                                             GENERAL INFORMATION:
JAPPIICANT: Schulein, Martin
APPIICANT: Schulein, Martin
APPIICANT: Schulein, Martin
JAPPIICANT: Biornvad, Mads Erik K.
APPIICANT: Biornvad, Mads E.
APPIICANT: Schnorr, Kirk
JITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-US
CURRENT PRILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: US/09/198,956
CURRENT FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-02
EARLIER FILING DATE: 1997-11-02
SARILER FILING DATE: 1997-11-02
SARILER FILING DATE: 1997-11-02
SOFTWARE: FASESEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH, 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 35.2%; Score 165; DB 3; Length 493; Best Local Similarity 36.1%; Pred. No. 4.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 WAAM-----GCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6187580el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/198,955A
CURRENT FILING DATE: 1998-11-24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION UNDBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR PLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1997-12-02
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-10-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 12, Application US/09198955A Patent No. 6187580 GENERAL INFORMATION:
Sequence 10, Application US/09198956
Patent No. 6165769
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Andersen, Lene N. Schulein, Martin Lange, Niels E. Bjornvad, Mads E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-198-956-10
```

```
US-10-072-152-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                           1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                           Query Match 35.2%; Score 165; DB 3; Length 493;
Best Local Similarity 36.1%; Pred. No. 4.1e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 35.2%; Score 165; DB 4; Length 493; Best Local Similarity 36.1%; Pred. No. 4.1e-10; Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps
                                                                                                                                                                                                                                                                                                                      390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: MOILET, Soren
APPLICANT: MOILET, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Schopen, Markus S.
APPLICANT: Schopen, Markus S.
APPLICANT: Schopen, Lars
TITLE OF INVENTION: NO. 6368843el Pectate Lyases
TITLE PREERENCE: 5378.200-US
CURRENT FILING DATE: 1200-10-23
PRIOR APPLICATION NUMBER: US/09/694,531
CURRENT FILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR PILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-02
PRIOR FILING DATE: 1997-12-02
PRIOR PILING DATE: 1997-11-02
PRIOR PILING DATE: 1997-12-02
PRIOR PILING DATE: 1999-11-02
PRIOR PILING DATE: 1999-11-02
NUMBER OF SEQ ID NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAM-----GCGNIRASFGSVNPATPTADIYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAM-----GCGNIRASFGSVNPATPTADTYLO 88
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-198-955A-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
US-09-670-141-10
; Sequence 10, Application US/09670141
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 12, Application US/09694531
Patent No. 6368843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Andersen, Lene N. APPLICANT: Schulein, Martin APPLICANT: Lange, Niels E. APPLICANT: Bjornvad, Mads E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-694-531-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-694-531-12
                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                         ઠે
```

```
SERREAL INCOMENTOR:
APPLICANT ANGERED.
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AGES ETHEN
APPLICANT SCHOOL AG
```

```
APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 TSSMSVEFYNSNKSAQTWSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         34; Indels 12; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; DB 1; Length 162;
4.2e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION UNDER: US/08/048,164A
FILING DATE: U4-APR-1993
CLASSIFICATION: 435
ATTONREY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-003
TELEPAN: (212) 790-9090
TELEPAN: (212) 790-9090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 AGALLGNSYVDNTSKVTANFVKETASP-TSTYDTYVE 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 AAMGCGN-----IRASF--GSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                        APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Joi, Roy H.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24.3%; Score 114; DB 29.9%; Pred. No. 4.2e-tive 22; Mismatches
                                                                   491 AQIGRINVLLAF -- ANFTGSNTDTY 513
                             86
                           62 AAMGCGNIRASFGSVNPATPTADTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                           Sequence 2, Application US/08048164A Patent No. 5496934 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/08460462 Patent No. 5670623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 24.39
Best Local Similarity 29.99
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: New IC...
STATE: New York
TOTAL U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                         US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-048-164A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-460-462-2
                           ઠે
                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor CITY: New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 SGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCDW
                                                                                                                                                                             10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 7, Application US/07862588B
Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
                                                                                                                              35.2%; Score 165; DB 4; Length 493; 36.1%; Pred. No. 4.1e-10; ive 23; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                      ||: :| ||||| 390 HAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
                                                                                                                                                                                                                                                                                                         61 WAAM------GCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .5e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25.7%; Score 120.5; I 32.9%; Pred. No. 3.5e-tive 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                          ; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-10-072-152-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME: Zelson, Steve T. / Lam
REGISTRATION NUMBER: 30,335
REFERENCE/DOCKET NUMBER: 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELECOMMUNICATION INFORMATION: TELEPHONE: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 531 amino acids TYPE: AMINO ACID
                                                                                                                                                  Best Local Similarity 36.18
Matches 35; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 25.7%
Best Local Similarity 32.9%
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       212 867 0298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 19920727
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 212 867 0296
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10017
                                                                                                                                                                                                                                                                                                                                                                                                                                         US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-07-862-588B-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
SEQ ID NO 12
LENGTH: 49
                                                                                                                                  Query Match
                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                         ò
```

```
CORRESPONDENCE ADDRESS:

ADDRESSEE PENNER & EDMONDS
STREET: 1155 Avenue of the Americas
COMPUTEN: New York
STATE: New York
STATE: New York
CONPUTEN: New York
STATE: New York
COMPUTEN: New York
COMPUTEN: New York
MEDIUM TYPE: Ploppy disk
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: Pacentin Release #1.0, Version #1.25
COMPUTEN: The PC-DOS/MS-DOS
SOFTWARE: The PC-DOS/MS-DOS
SOFTWARE: THE PC-DOS/MS-DOS
SOFTWARE: THE PC-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-DOS/MS-
```

3 TSSMSVEPYNSNKSAQTNSITPIIKITNTSDSDLNLNDVKVRYYYTSDGTQGQTFW-CDH 61

a

ò

62 AAMGCGN-----IRASF--GSVNPATPTADTYLO 88

۶.

```
1 VSGGVKVQYKANDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 WAAMGCGNIRASFGSVNPATPTADTYLO
US-09-917-376-5
190
180
165
165
165
165
150.5
                                                                                                LENGTH: 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE
 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Appli
Appli
                                                                         May 11, 2004, 12:09:36 ; Search time 10.2683 Seconds (without alignments) 2378.773 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 5,
Sequence 4,
Sequence 4,
Sequence 1,
Sequence 5,
Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Sequence 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1
Sequence 1
Sequence 6
                                                                                                                                          1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                              | Cgn2 6/ptodata/2/pubpaa/USO7 PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO7 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO6 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO6 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO6 PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/PCTUS PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USO8 NEW PUB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
| Cgn2 6/ptodata/2/pubpaa/USIOR PUBCOMB.pep:*
          GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5
US-10-155-400-5
US-09-1776-4
US-09-917-376-4
US-09-917-378-1
US-09-917-378-5
US-09-917-384-5
US-09-917-384-6
US-09-917-384-6
US-09-917-384-6
US-09-917-384-6
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-09-917-384-1
US-10-155-400-1
US-10-155-400-1
US-10-155-400-1
                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                  1140673 segs, 277566755 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                         Published Applications AA:*
                                                                                                                                                                                                                                                                           Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                     OM protein - protein search, using sw model
                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                             seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length DB
                                                                                                                    US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             150
1043
1043
1228
1228
1228
957
957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing:
                                                                                                                    Title:
Perfect score:
                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum DB
Maximum DB
                                                                                                                                            Sequence:
                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                         Database
                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Result
```

```
Sequence 29, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 51483, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 406, Applace and Applace 24, Appli Sequence 24, Appli Sequence 20, Appli Sequence 61, Appl Sequence 61, Appl Sequence 20, Appl Sequence 25, Appl Sequence 25, Appl Sequence 25, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 15, Appl Sequence 1577, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Sequence 1777, Appl Se
                                                                                                                                                                                                                                                                                                                  Sequence 223, App
Sequence 53, Appl
Sequence 53, Appl
Sequence 101, App
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: DING, SHI-YOU
APPLICANT: DING, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, TODD B.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.44501
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT APPLICATION NUMBER: US/09/917,376
SURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver: 2.1
SEQ ID NO 5
                                                                          Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
2 US-10-312-312-20
2 US-10-312-312-61
US-09-747-835A-25
2 US-10-312-312-25
2 US-10-424-599-196885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 469; DB 12;
100.0%; Pred. No. 1.5e-48;
trive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , OTHER INFORMATION: Carbohydrate binding domain US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-764-860-377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 5, Application US/09917376 Publication No. US20040038334A1 GENERAL INFORMATION:
         Best Local Similarity 100.
Matches 88; Conservative
                                                                            986
1131
1325
1325
1346
1346
1346
1371
```

```
61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/917,378
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 154
TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURE:
NAME/KEY: MOD RES
COCATION: (89)
COTHER INFORWATION: Any amino acid
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 4, Application US/09917378 Publication No. US20030119093A1 GENERAL INFORMATION:
                                                                                                                  Sequence 4, Application US/10155400 Publication No. US20030108988A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 100.
Matches 88; Conservative
                                                                                                                                                             GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          o;
                                                                                APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: TODB B.
APPLICANT: HIMMEL, MICHAEL B.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS FITLE OF INVENTION: CELLLOLOLYTICUS
FITLE OF INVENTION: CELLLOLOLYTICUS
FILLE REFERENCE: NREL 01.36A
CURRENT APPLICATION NUMBER: US/10/155,400
NUMBER OF SEQ ID NOS: 2002-10-22
NUMBER OF SEQ ID NOS: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ch 100.0%; Score 469; DB 14;
1 Similarity 100.0%; Pred. No. 1.5e-49;
88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 100.0%; Score 469; DB 12; Best Local Similarity 100.0%; Pred. No. 1.5e-48; Matches 88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                              ; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-10-155-400-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ) OTHER INFORMATION: Any amino acid

) OTHER INFORMATION: Carbohydrate binding domain

US-09-917-376-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILE REFERENCE: 40197.4US01
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN VET: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM: Acidothermus cellulolyticus
                   Sequence 5, Application US/10155400
Publication No. US20030108988A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 4, Application US/09917376
Publication No. US20040038334A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 4
LENGTH: 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION:
                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate
OTHER INFORMATION: binding domain
US-08-917-378-4
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHAEL DIERANT AVICELASE FROM ACIDOTHERMUS
TITLE OF INVENTION: TELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
TITLE OF INVENTION: CELLULOLYTICUS
CURRENT APPLICATION NUMBER: US/10/155,400
CURRENT FILING DATE: 2002-10-22
NUMBER OF SEO ID NOS: 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: ADDER, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL, MICHAEL B.
AITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.7US01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Score 469; DB 14;
; Pred. No. 1.5e-48;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 100.0%; Score 469; DB 10; Best Local Similarity 100.0%; Pred. No. 2.9e-48; Matches 88; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; OTHER INFORMATION: Carbohydrate binding domain US-10-155-400-4
```

ö

```
Pred. No. 6.5e-48;
1; Mismatches 0;
                                                                                                                                                        61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 WAAMGCGNIRASFGSVNPATPTADTYLQ 88
                                                                                                                                                                                                                                                                                               US-09-917-383-5
; Sequence 5, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
  98.98;
Best Local Similarity 98.9
Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 87; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            i Other AMPONT
US-09-917-383-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                             요
                                                                                                                                                          ò
                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                   9
                                   455 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 60
       1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                       Sequence 1, Application US/09917378

Publication No. US20030119093A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: DING, SHI-YOU

APPLICANT: VINZANT, TODD B.

APPLICANT: VINZANT, TODD B.

APPLICANT: HIMMEL, MICHAEL B.

TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

TITLE OF INVENTION: CELLULOLYTICUS

SOFTWARE: DALOCHTON

SOFTWARE: PALENTING

SEQ ID NO 1

LENGTH: 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 5, Application US/09917384

Publication No. US20030096342a1

GENERAL INFORMATION:

APPLICANT: DING, SH-YOU

APPLICANT: DING, SH-YOU

APPLICANT: DING, SH-YOU

APPLICANT: HIMMEL, MICHER B.

APPLICANT: HIMMEL, MICHER B.

TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 40170.64021

CURRENT FILLING DATE: 201-07-28

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PATENTIN Ver. 2.1

SEQ ID NOS: 4

LENGTH: 150

TYPE: PRT

CREANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 10; Length 762;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99.4%; Score 466; DB 10; Length 150;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 100.0%; Score 469; DB 10; Length Best Local Similarity 100.0%; Pred. No. 2e-47; Matches 88; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION: Description of Artificial Sequence: OTHER INFORMATION: GuxA
                                                                                                 61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                              61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-384-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE:
                                                 셤
                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                       9
                                                                                                                                                      1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINZANT, TODD B.
APPLICANT: VINZANT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40170.6USQ1
CURRENT APPLICATION NUMBER: US/09/917,383
CURRENT APPLICATION NUMBER: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SOFTWARE: PATENTIN OF SEQ ID NOS: 17
CURRENT APPLICATION OF SEQ ID NOS: 17
SOFTWARE: PATENTIN OF SEQ ID NOS: 17
CURRENT SEQ ID NOS: 17
CURRENT SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN OF SEQ ID NOS: 18
CONTANTANTES PATENTIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 6, Application US/09917384

Publication No. US20030096342A1

GENERAL INFORMATION:

APPLICANT: DING, SHILIAM S.

APPLICANT: ADNEY, WILLIAM S.

APPLICANT: HIRWEL, MICHALE B.

TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REFERENCE: 40170-61991

CURRENT APPLICATION NUMBER: US/09/917,384

CURRENT PILING DATE: 2001-07-28

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PATENTIN Ver. 2.1

SEQ ID NO 6

LENGTH: 1043

TYPE: PRT

ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 150,
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99.4%; Score 466; DB 10;
98.9%; Pred. No. 6.5e-48;
tive 1; Mismatches 0;
```

```
US-09-917-383-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-917-383-1
                                                                                             FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                       69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGCVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/09917384

Publication No. US20030096342A1

GENERAL INFORMATION:
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: HIWMEN, TODD B.
APPLICANT: HIWMEN, TODD B.
APPLICANT: HIWMEN, THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REPERRINCE: 40170 6US01
CURRENT APPLICATION NUMBER: US/09/917,384
CURRENT FILING DATE: 20010-07-28
NUMBER OF SEQ ID NOS: 14
                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 6, Application US/09917383
; Sequence 6, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION
; APPLICANT: DING, SHI-YOU
; APPLICANT: VINZANT, TOOD B.
; APPLICANT: VINZANT, TOOD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION CELLULOLYTICUS
; TITLE OF INVENTION CELLULOLYTICUS
; TITLE OF INVENTION CELLULOLYTICUS
; TITLE OF INVENTION UNBER: US/09/917,383
; CURRENT FALING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; OTHER INFORMATION: Description of Artificial Sequence: Segment of ; OTHER INFORMATION: GuxA
US-09-917-384-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; OTHER INFORMATION: Description of Artificial Sequence: Segment of ; OTHER INFORMATION: GuxA
US-09-917-383-6
                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99.4%; Score 466; DB 10; Length 1043; 98.9%; Pred. No. 6.8e-47; ive 1; Mismatches 0; Indels 0
                                                                                             Length 1043;
                                                                                         Score 466; DB 10; Length 1
Pred. No. 6.8e-47;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                             537 WAAMGCGNIRASFGSVNPATPTADTYLQ 564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                           Query Match
Best Local Similarity 98.9%;
Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AOFTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 98.9 Matches 87; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH: 1043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
US-09-917-384-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              엄
                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                ð
```

```
584 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRVRYFTRDGGSSTLVYVCD 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: DING; SHI-YOU
APPLICANT: DING; SHI-YOU
APPLICANT: DING; SHI-YOU
APPLICANT: DINGAN; TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: DECKER, STEPHEN R.
TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
TITLE OF INVENTION: THERMAL TOLERANT
TITLE OF INVENTION: CELLULALITICUS
TITLE REFERENCE: 40170.6US01
CURRENT PLILOR DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 14
SSOTID NO 1
LENGTH: 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/09917376
Publication No. US20440038334A1
GENERAL INFORMATION
APPLICANT: DING, SHI-YOU
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: CHELULOLYTICUS
FILE REFERENCE: 40197.4US01
                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Segment
OTHER INFORMATION: GuxA
                                                                         OTHER INFORMATION: Description of Artificial Sequence: Segment OTHER INFORMATION: GuxA
US-09-917-384-1
                                                                                                                                                                                           Length 1228;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1228;
                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 99.4%; Score 466; DB 10;
Best Local Similarity 98.9%; Pred. No. 8.3e-47;
Matches 87; Conservative 1; Mismatches 0;
                                                                                                                                                                                        Query Match 99.4%; Score 466; DB 10; Best Local Similarity 98.9%; Pred. No. 8.3e-47; Matches 87; Conservative 1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 644 WAAMGCGNIRASFGSVNPATPTADTYLQ 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1, Application US/09917383; Publication No. US20030104522A1; GENERAL INFORMATION:
TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: Artificial Sequence
```

```
ે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       원
                                                                                                                                                                                                                                                                                                                                                                                                                                          869 VSGGVKVQYKANDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          869 VSGGVKVQYKANDSAPGDNQIKPGLQVVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD 928
                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                  1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 VSGGVKVQYKNNDSAPGDNQIKPGLQLVNTGSSSVDLSTVTVRYWFTRDGGSSTLVYNCD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: DING, SHI-YOU
APPLICANT: DING, SHI-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: TOTAGAT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: NREL 01-36A
CURRENT APPLICATION NUMBER: US/10/155,400
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PATENTIN Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ch 98.5%; Score 462; DB 14; Length 957; 1 Similarity 97.7%; Pred. No. 1.9e-46; 86; Conservative 2; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                      Query Match 98.5%; Score 462; DB 12; Length 957; Best Local Similarity 97.7%; Pred. No. 1.9e-46; Matches 86; Conservative 2; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 WAAMGCGNIRASFGSVNPATPTADTYLO 88
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SEQ ID NO 1
LENGTH: 957
                                                                                                                                      TYPE: PRT
ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 8100, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OWURA, SATOSHI
APPLICANT: IKEDA, HARUO
                                                                                                                                                                                  FEATURE:

NAME/KEY: MOD RES
LOCATION: (957)

CTHER INFORMATION: Any amino acid
US-09-917-376-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1, Application US/10155400 Publication No. US20030108988A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; NAME/KEY: MOD_RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 86; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-156-761-8100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH: 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-155-400-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

```
| APPLICANT: ISHIKAWA, JUUN
| APPLICANT: HORIKAWA, HIROSHI
| APPLICANT: SHIBA, TADAYOSHI
| APPLICANT: SHIBA, TADAYOSHI
| APPLICANT: SHIBA, TADAYOSHI
| APPLICANT: SHIBA, TADAYOSHI
| APPLICANT: SHIBA, TADAYOSHI
| APPLICANT: MASAHIRA
| TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
| FILE REFERENCE: 149-26.
| CURRENT APPLICATION NUMBER: US 2001-05-29
| PRIOR FILING DATE: 2002-05-29
| PRIOR PELICATION NUMBER: US 2001-204089
| PRIOR FILING DATE: 2001-05-09
| PRIOR FILING DATE: 2001-08-02
| NUMBER OF SEQ ID NOS: 15109
| PRIOR FILING DATE: 2001-08-02
| NUMBER OF SEQ ID NOS: 15109
| LENGTH: 741
| TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: PRT | TYPE: TYPE: PRT | TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYPE: TYP
```

Run on:

```
Z86105 A. thermophi
AF078038 Caldicell
AF078038 Caldicell
AF355629 Bacillus
Z29076 Bacillus su
MIG185 B. subtilis
X67044 B. subtilis
X04889 Bacillus su
AY04252 Bacillus
Z73234 B. subtilis
Z99113 Bacillus
Z99113 Bacillus
Z99113 Bacillus
Z9913 Bacillus
AX183475 Bacillus
AX183475 Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B05425 DNA sequenc
AB016164 Bacillus
A28470 B.lautus st
X13602 Caldocellum
U16308 Caldocellum
M76588 Bacillus la
U27084 Bacillus sp
                                                                                                                                                                                                                                                                                                                             AY298814 Thermobif
AF163837 Caldibaci
AF200304 Caldibaci
AJ488933 Paenibaci
AJ133614 Bacillus
E35142 Truncated C
                                                                                                                                                                                                                                                                    AX700050 Sequence
AX700036 Sequence
AX700058 Sequence
AX700025 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF363635 Bacillus
AF155197 Acetivibr
AL939105 Streptomy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            L32742 Caldocellum
M28332 B.subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   L01257 Caldocellum
                                                                                                                                                                                                                                                                                                                                                                                                             AF078042 Caldicell
E35100 Truncated c
                                                                                                                                                                                                                                                                                                                                                                                                                                      AB008029 Thermoph
AF078737 Caldicel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF078040 Caldicel
                                                                                                                                                                                                                                                                                                                                                                                                                                                             E35099 Truncated
                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                AF078737
E35099
AF078038S3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATZ86105
AF078038S1
CASR69XYN2
                                                                                                                                                                                                                                                                                                                                                                                                    E35143
AF078038S5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BACCMCASE
AY183475
AF045482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BSBGLUC2
AY044252
BC170DEGR
BSUB0010
                                                                                                                                                                                                                                                                                                                                                      AF200304
PAE488933
BSP133614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF355629
BSEGLSG
BACGLUB
BS14GLUC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E05425
AB016164
A28170
                                                                                                                                                                                                                                                                                                                                                                                                                                      AB008029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CDCMANA
CDCCELA
BACCELD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SC093910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2SU16308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AF15519'
                                                                                                                                           em_htgo_other:
                      em_htg_inv:*
em_htg_other:*
                                                                                                                              htgo_mus:*
                                                                                   em_htg_mam:*
em_htg_vrt:*
em_sv:*
                                                                                                                  htgo_hum:*
          em_htg_hum:*
em_htg_inv:*
                                                         em_htg_pln:*
em_htg_rod:*
                                                em_htg_mus:*
                                                                                                                                                                                                                          %
Query
Match Length DB
em_vi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5513
3262
5437
1523
1800
1920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2589
26170
207829
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5439
2175
2175
2587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1593
2084
4977
                                                                                                           87.
                                                                                                                    e e
                                                                                                          ea
E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             180
177
172.5
168
                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
186
186
182.5
182.5
182.5
                                                                                                                                                                                                                                                                                                                                                        214.5
209.5
209.5
195.5
194.5
192.5
192.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     191
190.5
190.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    186.5
186.5
                                                                                                                                                                                                                                                                                                                                                                              Result
                                                                                                                                                                                                                                                  ġ
                                                                                                                                                                                                                                                                                                                                                                         Command line parameters:
-MODEL=frame+ph.model_DBV=xlh
-MODEL=frame+ph.model_DBV=xlh
-MODEL=frame+ph.model_DBV=xlh
-DB=GanEmbl -QFWT=ph.model_DBV=xlh
-DB=GanEmbl -QFWT=fastap_SIDFIX=pln.rge -MIRMATGH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=buman40.cdi -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=buman40.cdi -LIST=45
-OOCALIGN=200 -THR_SCORE=pct -THR MAX=100 -TRR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINIEN=0 -MAXIEN=200000000
-USER=USO9917376 @CGN 1 1 7742 @runat 11052004 114535 28241 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -MAIT -DSPBELOKE=100 -LONGLOG
-DBV TIMEOUT=120 -WARN TIMEOUT=30 -THERADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
                                                                                            (without alignments)
3015.368 Million cell updates/sec
                                                                                                                                                    1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                May 12, 2004, 03:40:05 ; Search time 1264.92 Seconds
                                                                                                                                                                                                                                                                        6940544
            GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                         - nucleic search, using frame plus p2n model
                                                                                                                                                                                                                                               3470272 seqs, 21671516995 residues
                                                                                                                                                                                                                                                                      hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                        BLOSUM62
Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                              US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gb_htg.;
gb_ov.;
gb_ov.;
gb_pt.;
gb_pt.;
gb_rc.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gb_sy:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
em_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_pat: *
em_ph: *
em_pl: *
em_ro: *
em_sts: *
em_ats: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_om:*
em_or:*
em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_mn:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GenEmbl:*
                                                                                                                                                                                                                                                                      Total number of
                                                                                                                            Title:
Perfect score:
Sequence:
                                                                                                                                                                            Scoring table:
                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                                                 Searched:
```

Streptom

```
1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                              1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC 1929
                                                                                                                                                                                                                                                                                                                                                             1930 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ValSerGlyGlyValLysValGlnTyrLysBsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                      41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                              TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2605 GTGTCGGGTGGGGTGAAAGTGCAGTATAAGAATAATGATTCGGCGCCGGGTGATAATCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Acidothermus cellulolyticus
Acidothermus cellulolyticus
Bactoria; Actinobacteria; Actinobacteridae; Actinomycetales;
Frankineae; Acidothermaceae; Acidothermus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding, S.Y., Adney, W.S., Vinzant, T.B. and Himmel, M.E.
Thermal tolerant avicelase from Acidothermus cellulolyticus
Patent: WO 033012090-A 2 13-FEB-2003;
Midwest Research Institute (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear

    .2869
    ^organism="Acidothermus cellulolyticus"
/mol type="unassigned DNA"
    /db_xref="taxon:28049"

                                         3687
87
1
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2869
86
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AX700058 2 from Patent W003012090.
AX700058
                                                                                                    Indels:
                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-5 (1-88) x AX700036 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x AX700058 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AX700058.1 GI:29536021
                                          9.53e-48
466.00
100.00%
98.86%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2.19e-47
462.00
100.00%
97.73%
98.51%
                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                   1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
AX700058
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                g
                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    심
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAT 03-APR-2003
                PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1483 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGGTCGTCGTCGACTGGTACAACTGTGAC 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ding,S.Y., Adney,W.S., Vinzant,T.B., Himmel,M.E. and Decker,S.R. Thermal tolerant cellulase from Acidothermus cellulolyticus Patent: WO 03012109-A 2 13-FEB-2003; Midwest Research Institute (US)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1423 ATCAAACCGGGTCTCCAGTTGGTGAATACGGGGTCGTCGTCGTGGATTTGTCGACGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                 Acidothermus cellulolyticus
Acidothermus cellulolyticus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                       Ding, S.Y., Adney, W.S., Vinzant, T.B. and Himmel, M.E. Thermal tolerant mannanaee from Acidothermus cellulolyticus Patent: WO 03012110-A 2 13-FEB-2003; Midwest Research Institute (US) Location/Qualifiers
                                                                                                                                                                                                                                           /organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="taxon:28049"
                                                                                                                                                                                                                                                                                                                                2289
88
0
0
0
                                                                                                                                                                                                                                                                                                                                                            Conservative:
Mismatches:
Indels:
                                                                                                                            Frankineae, Acidothermaceae, Acidothermus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism="synthetic construct"
/mol_type="unassigned DNA"
db xref="taxon:32630"
/noTe="Segment of GuxA"
                DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                                                                Length:
Matches:
        AX700050 2289 bp
Sequence 2 from Patent W003012110.
AX700050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AX700036
Sequence 2 from Patent W003012109-
AX700036 AX700036.1 GI:29536019
                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x AX700050 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                        AX700050.1 GI:29536020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  artificial sequences.
                                                                                                                                                                                                                                                                                                                             2.21e-48
469.00
100.00$
100.00$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    synthetic construct
synthetic construct
                                                                                                                                                                                                                               1. .2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1. .3687
                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
AX700036
LOCUS
DEFINITION
                         DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                              No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                         REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEYWORDS
SOURCE
XX700050
                                                                                                                                                                                                                PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PEATURES
                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
12 (Dases 1 to 299175)

Somura,S., Ikeda,H., Ishikawa,J., Hanamoto,A., Takahashi,C.,
Shinose,M., Takahashi,Y., Horikawa,H., Nakazawa,H., Osonoe,T.,
Kushida,N., Director-General of Biotechnology Center, Shiba,T.,
Sakaki,Y. and Hartori,M.
Direct Submission

Direct Submission

Submitted (29-MAR-2002) Director-General of Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center; 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: biodenite-go.jp, WRL:http://www.bio.nite-go.jp/,
Tel:81-3-3481-1933, Fax:81-3-4842,
This work was done in collaboration with Haruo Ikeda(*1), Jun Ishikawa(*2), Akiharu Hanamoto(*3), Ghiguas Takahashi(*3), Mayumi Shinose(*3), Hiroshi Horikawa(*4), Hidekazu Nakazawa(*4), Tadayoshi Shiba(*5), Yoshiyuki Sakaki(*6,*7), Masahira Hattori(*1,*7)
and Satoshi Omura(*1,*3).
Final finishing process and all annotation were done by H. Ikeda and J. Ishikawa.
                                                                                                                                                                                                                                                                                                                                             Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis
Nat. Biotechnol. 21 (5), 526-531 (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /translation="MAGPPRSATRSRTTADREDSAHPAPDRADEPGYRRQIKVQANLQ
EGHHALARRAGQLYQRYQDGMEDQIGALGLVLNALVLFNTRYMDAALTQLRADGFEVR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Streptomyces avermitilis MA-4680"
/mol type="genomic DNA"
/strāin="MA-4680"
/db xaref="taxon:227882"
/nofe="This strain is also named as strain: ATCC 31267,
NCIMB 12804 or NRRL 8165.~synonym: Streptomyces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       *1 Kitasato Institute for Life Sciences, Kitasato University
*2 National Institute of Infectious Diseases
                                                                                                                                               Shinose, M., Takahashi Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y. and Hattori, M. Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary
                      Streptomyces avermitilis MA-4680
Streptomyces avermitilis MA-4680
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                               Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C.
                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 98 (21), 12215-12220 (2001)
21477403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                *3 The Kitasato Institute
*4 National Institute of Technology and Evaluation
*5 School of Science, Kitasato University
*6 Institute of Medical Science, University of Tokyo
*7 RIKEN, Genomic Sciences Center
                                                                                       Streptomycineae; Streptomycetaceae; Streptomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              codon start=1
/transI_table=11
product="putative transposase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEDVARLSPFMRHHINVLGRYSFOLPDLE"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note="IS630 family, truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Following url is also available.
http://avermitilis.ls.kitasato-u.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /protein_id="BAC68155.1"
/db_xref="G1:29604084"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    complement (155. .550)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tn4561, truncated"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                avermectinius
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="SAV445
                                                                                                                                                                                                                                               metabolites
                                                                                                                                                                                                                                                                                                           11572948
                                                                                                                                                                                                                                                                                                                                                                                                                                                          22608306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                                                                                                                                                                                                                                              JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
MEDLINE
                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                           PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PUBMED
  KEYWORDS
                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                              TITLE
                                                                                                                                                                                                                                                                                                                                                                                              TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMMENT
                                                                                                           PAT 03-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BCT 10-MAY-2003
                                          2725 ACGSTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       112 CTCAAAGCGCAGTATAAGAACAATGATTCGGCGCCGAGTGACAACCAGATCAAACCGGGT 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        292 GGGTGTGGGGGATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACGCCGACGGCGGAC 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 232 TGGTTCACCCGGGATGGTGGTCGTCGACACTGTGTACAACTGTGACTGGGCGGCGATG 291
                                                                                       80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                and
  9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           84
ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                    TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Thermal tolerant exoglucanase from Acidothermus cellulolyticus Patent: WO 03012095-A 2 13.FEB-2003; Midwest Research Institute (US) Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AP005023 299175 bp DNA linear BCT 10-MAY.
Streptomyces avermitilis genomic DNA, complete genome, section
                                                                                                                                                                                                                                                                                                                                                                                                                Acidothermus cellulolyticus
Acidothermus cellulolyticus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Adney, W.S., Ding, S.Y., Vinzant, T.B., Himmel, M.E.,

    .3365
/organism="Acidothermus cellulolyticus"
/mol_type="unassigned DNA"
/db_xref="t-axon:28049"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Frankineae; Acidothermaceae; Acidothermus
                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                                                                                                                                                                  Sequence 2 from Patent WO03012095.
AX700025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                  2845 ccgaccccacciraccrccac 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-5 (1-88) x AX700025 (1-3365)
                                                                                                                                                                            88
                                                                                                                                                                            ProThrAlaAspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3/30.
AP005023 BA000030
AP005023.1 GI:29604083
                                                                                                                                                                                                                                                                                                                                                                          AX700025.1 GI:29536018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.72e-44
436.00
97.62%
96.43%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lantz Mccarter, S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                       61
                                                                                                                                                                            81
  41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
AP005023
                                                                                                                                                                                                                                                                                   AX700025
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                      a
                                                                                       ò
                                                                                                                        g
                                                                                                                                                                         ò
                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
105341 TCGGGCGCCTCAAGGTCCTCTACAAGAACAACGACTCCTCGGCCACCGACAACGAC 105400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               105461 GCCGGCTACTACTTCACCGGGGACAGCGGCTCGCCCACCGTGAAGGCCTGGTGCGACTAC 105520
                                                                                                                                                                                                                                  /translation="mytayrararagsawqacrgarrssqhssegyrrmdekeflrty
aertglsrqeaadltwatletlarrlspgeardlitelpeglaeavrrgttdrierfg
hhdsyrrvaertwlkeeeadrgyrtylaylreaisekefndlmsqlgtdfsqalesag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVGARVRWQVDTWTAPLSVSEQADVSDIAEAVRSHLIGGAEWDVAIETLTDFPRRARLYP
ISVEVDTALRAALISLPALGYRRLRRRVRQAVVDVVRELVAHEETPRHPEVIGRPPRE
PVPDTROEPEQEQRYVVPGLRGHTRLVAGMYTANRPWRLRGSLSRALAGVFATATVLF
INATWINLATAGAVQHAYASSAVITWII IVDHQLWERGTHLPARHPLYPYNLV
TLVTIALAVVVLAAALFATLVAVSFLLLIDASVIRKFIGRQVGVGDYLFLAWFITAIAM
IGGAFGTSLEGDEKVSDAAYGRRQRREQRMLRQMPRPERGERRGPDTTDG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /translation="mmpyarasgieorpgwswidbylmcscdiiadiabgrieorplu
atvarleagervlavgpaorvtwralgdgrymhonavargsparvlgsmasmatsalg
nsarrraaandaoprcgartpglhvidvldvhgtmlrmsmhigtsstrrpssalrall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /translation="MATLLRRVCPHEGURIAAVPAGFVLSRRPWTASVYRCAVVGPQS
SEREBEWGGSRTPVAAASSRLAPCRRRLSVVRCPRRAVPSQAVRSRFRPCTTRFVGDVL
VVQRRDALHARVPHVIDDVHPGCRGAARPVRIATHSLGGVIAVDMATANVPLWTES
LVTFGSQATFFHQPVFGDLLCQMIAFPMRKHQBAETGALPPSTGF1AHAGQGW
SAPRITDIEDPEVHRTLRRSLEDLGAPAHLRPAAPRTHGGRRRCPLVHGPVRPRLAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              translation="MIMTADVPETSSGRSVRVALVADPGLPSEIARALAPRLPDHLRR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValargTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 449
113
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
| transl_table=11
| product="hypothetical protein"
| protein_id="BAC68154.1"
| db_xref="GI:29604093"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1

{rransl_table=11

{producf="hypothetical protein"

{protein_id="BAC68165.1"

{db_xref="GI:29604094"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start=1
|transl_table=11
|product="hypothetical protein"
|protein_id="BAC68166.1"
|db_xref="GI:29604095"
                                                                                                                                                      /product="hypothetical protein"
/protein_id="BAC68163.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             complement (6986. .7498)
/note="SAV455"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x AP005023 (1-299175)
                                                                                                                                                                                                          db xref="GI:29604092"
                                                                                                       codon_start=1/transl_table=11
                              LVHGTAAGPCTPR"
                                                                                                                                                                                                                                                                                                                                               5810. .6964
/note="SAV454"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7548. .8189
/note="SAV456"
                                                                         /note="SAV453"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.91e-19
257.00
71.26%
56.32%
54.80%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ELVDLHVIAL"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                 /transl_table=11
/product="putative transposase"
/protein_id="BAC68157.1"
/db_xref="G1:2564086"
/translation="WVVVAVGCALTSVPTAGGVVAEPVCVRRLTDQEGQKLQQIVRRG
SASSVRFRRAMMLLASAGGRRVPVIAQLV"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /codon statt=1
/transl_table=11
/product="putative TetR-family transcriptional regulator"
/product="putative TetR-family transcriptional regulator"
/protein id="BAGG8158.1"
/db_xref="G1:29604087"
/translation="MPTSTRRARERAWTRERIIEAALHVLETEGIAALTIRRIATDVE
YSAPVVYGHRANKDALVLELVAHGHRAMLSEFSQAAQEPDTDRRWTRIASQYVRFAGE
HPHLEVOWNUPRWYDBRUNDERRRYAREPVIDVLKELLTSWAAHDVVLADFDQACEILWGTL
YGIASLSYLGSVGNDRARYLABQALHTILLGWRTETPANGRPASS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /transl_table=11
/product="hypothetical protein"
/protein id="BAC68167.1"
/db_xref="G1:29664091"
/translation="WRAVKESRVRVVGEIAATPTGEALELVLHLGESLRGVRVESSDA
TARPDGLDFLAELPARGEWHTTALVRPSVEGEEAGGLFHPGAEPGESGPARRLCAWRE
                     /codon_start=1
/transl_table=11
/product="gutative transposase"
/protein_id="BAC68156.1"
/protein_id="BAC68156.1"
/bc.xref="gutative transposase"
/translation="WPDTQNAGPSTSVTAAAVSTAGSVYSWMAPTSPVMAPPSPGRRR
EPGSLGAVRAYDVIRMLSRDGRPTPLGDAITYYGGSRRLCTSCA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon start=1
/transl_table=11
/product="hypothetical protein"
/product="hypothetical protein"
/product="hypothetical protein"
/db_xref="c1:25604088"
/db_xref="c1:25604088"
/db_xref="c1:25604088"
/db_xref="c1:25604088"
/db_xref="c1:25604088"
/db_yref="c1:25604088"
/db_yref="c1:25604088"
/db_yref="c1:25604088"
/db_yref="c1:25604088"
/db_yref="c1:2560408"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation="MAKHVIIGAGSIGTNVARLLABGGESVRIVTRSGSGPEHPLVDR
VADABASDPSRLTELSRGAGVIYHCANPPSYTVMERLLPPLGTAAITAAKANDAVLALT
GSLYAYGPQPGGRMNEHTPMAATGHKRELRRWWEQALAGGIRTVEVRGSDYIGKDAN
GIYSLLIKSALQKGRAAMITGHLDMHTFTFRGDMAQALITLASEBRAMGRPWHVPSA
PAVTIRELARRYAVAAGRPPIKLIQIPRSVTRTAGLIVPIAREMAEMDYQWYAPFHMD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /translation="WTDAFHIDVPHTDAALAVIALSGEFDIAAAPAVRARALELIANG
HPDLVADLSGVTFCDSSGLGALVGIWRCAKDADGSLTLAAIPDRLSRLLSVTGMDAFL
PAYSSADAALAARQGNRTTA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     producE="putative anti-anti-sigma factor"
protein id="BAC68161.1"
/db_xref="G1:29604090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /codon_start=1
trans1_table=11
product="hypothetical protein"
protein_id="BAC68160.1"
db_xref="GI:29604089"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATETTETFGLTATDLDTAVRNEVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  complement (4130. .4498)
/gene="bldG"
                                                                                                                                                                                                                                                                                                                                                                                                                                               complement (1954. .2571)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      complement (4130. .4498)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      codon_start=1
transT_table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           note="SAV451"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2694. .3098
/note="SAV449"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            127. .4029
note="SAV450"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      635. .5156
note="SAV452"
                                                                                                                                                                                                          766. .987
'note="SAV447"
                                                                                                                                                                                                                                                             codon start=1 trans table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             note="SAV448"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          codon_start=1
/transl_table=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /gene="bldg"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
```

```
/codon start=1
/transl_table=11
/product="unknown"
/protein_id="AAF22273.1"
/db_xref="G1:6651326"
/translation="TPTPTPSAAPTPTPSAGGSLVVQYRAADTNAGDNQLKPHFRI
VNRGTSSVPLSELTIRYWYTVDGDKPQVFNODWAQVGCSNVRGSFVKLSTGADYY
IEITFTSGAGSLAAGASSGDIQVRINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AF163837 1inear BCT 08-FEB-2000 Caldibacillus cellulovorans multidomain beta-1,4-mannanase precursor (manA) gene, complete cds; and unknown genes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                          2 (bases 1 to 4567)
Sunna,A., Gibbs,M.D. and Bergquist,P.L.
Direct Submission
Submitted (29-JUN-1999) Biological Sciences, Macquarie University,
North Ryde, New South Wales 2109, Australia
                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 4567)
Sunna, A., Gibbs, M.D., Chin, C.W., Nelson, P.J. and Bergquist, P.L.
A gene encoding a novel multidomain beta-1,4-mannanase from Caldibacillus cellulovorans and action of the recombinant enzyme
                                                                                                                                                                                                                                                                                                                   62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                               SerglyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                        42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61. .510
/note="Region: cellulose-binding domain type IIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Caldibacillus cellulovorans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae;
Caldibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . . 60
'note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Caldibacillus cellulovorans"
/mol_type="genomic DNA"
/db_xref="taxon:74586"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    kraft pulp
Appl. Environ. Microbiol. 66 (2), 664-670 (2000)
20120520
  N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /evidence=not_experimental
620. 3412
/gene="manA"
                                          US-09-917-376-5 (1-88) x AY298814 (1-1957)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Caldibacillus cellulovorans
                                                                                                                                                                                                                                                                                                                                                                                                                                       1748 GGCGCCGACACCTCCCTG 1765
                                                                                                                                                                                                                                                                                                                                                                                                   82 ThrAlaAspThrTyrLeu 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AF163837.1 GI:6651325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="orf1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NGQLIWGVEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10653733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF163837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
    DB:
                                                                                                                                                            ð
                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
/ Crainilation="MTPLTRRLRAGAAAIAIGASALIPLTSSPAAASGTADMLHTDGN
RIVDSAGNEWWITGANWFGETYSERNEHGLWAAUIBDITSAAABAEGINMVRVPISTQL
ILLEWRAGGAGPSGVNEYVNPELAGMYALLEDYAALDIKNEHDDHASAABABGINMVRVPISTQL
ILLEWRAGGAGPSGVNEYVNPELAGMYTLERYKNNDTIVAADIKNEFBIGKANETPRAARDGST
HYYPWXGGDITTEDPYTAWEWYTERYKNNDTIVAADIKNEFBIGKANETPRAARDGST
BYDPREHYGTAGERGITALLCEGIBIYPKOGGOWSSTDGRDYYSTWWGGNLA
GVADHPUDLGAHQDQLVYSPHDYGESVFROPEDPFEGEWNRQTLTEDWRRPNHYTHEDD
IAPLLIGEWGGSPLOGONBEKWHTALRSLIIDBKWHHTFWALLNRNSGDTGGLLAYDWTT
BPPTEPTERPALMPRGALEVYTRNNSLAADDSQIAPGIRLWYTGGGOPSGPPEPT
BPPTEPTEPTEPTEPPALWFGALEVYRNNSLAADDSQIAPGIRLWYTGSSTVDLAD
VEIHYYFTNEFGGTLQPTGOWAQWGCANVNASFTSLSAAGADTSLUTLLSGSLAPGAB
TELQGRIHTANWANFDESDDYSRGTNTDWELSEVITAYLGGTLWWGTPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 (bases 1 to 1957)
Posta,K., Beki,E., Kukolya,J. and Hornok,L.
Direct Submission
Submitted (14-MAY-2003) Agricultural Biotechnology and
Microbiology, Szent Istvan University Godollo, Pater Karoly str.1.,
Godollo 2103, Hungary
Location/Qualifiers
                                                                                                                                                                               AY298814 1957 bp DNA linear BCT 23-JUN-2003
Thermobifida fusca strain TM51 endoglucanase (ce15B) gene, complete
                                                                                                                                                                                                                                                                                                                             Bacteria; Actinobacteria; Actinobacteridae; Streptosporangineae; Nocardiopsaceae; Thermobifida.

1 (bases 1 to 1957)
Posta, K., Beki, E., Kukolya, J. and Hornok, L.
Phylogenetic relationships of If cel5B, a new endoglucanase encoding gene from Thermobifida fusca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cellulose binding domain; CBM-3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="Region: catalytic domain; GH5 family'
1532. .1765
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1957
45
16
22
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. 1957
/organism="Thermobifida fusca"
/mol_type="genomic DNA"
/strain="TM51"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /product="endoglucanase"
/protein_id="AAP56348.1"
/db_xref="GI:31745733"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db_xref="taxon:2021"
26._.39
/note="putative"
                                                                                               105581 GGAGCCGACCCTACCTCGAA 105601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             bound moiety="CelR" 07. .1957
                                                           82 ThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             codon start=1
trans1 table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  note="Region:
                                                                                                                                                                                                                                                            GI:31745732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /gene="cel5B"
07. 1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .07. .1957
gene="cel5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene="ce15B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene="cel5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.89e-18
228.50
70.93%
52.33%
48.72%
                                                                                                                                                                                                                                                                                                 Thermobifida fusca
                                                                                                                                                                                                                                                                                                                 Thermobifida fusca
                                                                                                                                                                                                                                                            AY298814.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                      AY298814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                      105521
                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . No. :
                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE .
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gene
                                                                                                                                      RESULT 6
AY298814
LOCUS
                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                           8
```

```
AF200304

3237 bp DNA linear BCT 01-DEC-2000 Caldibacillus cellulovorans putative type IIIb cellulose-binding domain gene, partial cds; and beta-1,4-xylanase XynA precursor (xynA) gene, complete cds.

AF200304.1 GI:7385018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /evidence=not_experimental
/transl_table=11
/product="putative type IIIb cellulose-binding domain"
/protein_id="AAF61648.1"
/bx.rotein_id="AAF61018"
/translation="YQCGSNIAGSFVKLSTGRTGADYYIEITFTSGAGSLAPGASSGD
                                                                                                                                                                                                     |||| ::: |||||||||||::|||
GGCAGCCTGGTCGTGCAGTATCGCGGGGGGGGAACGCGGGGGGACAACCAGCTGAAG 117
                                                                                                                                                                                                                                                                                                                                             235 CAGGTGGTTGTTCGAACGTGCGCGCAGCTTCGTGAAGCTTTCGACGGCCGGACGGG 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (29-007-1999) Biological Sciences, Macquarie University, North Ryde, Sydney, New South Wales 2109, Australia Location/Qualifiers
                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                   63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                      3 GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys 22
                                                                                                                                                                                                                                                                                                         1 (bases 1 to 3237)
Sunna,A., Gibbs,M.D. and Bergquist,P.L.
A novel thermostable multidomain 1,4-beta-xylanase from
'Caldibacillus cellulovorans' and effect of its xylan-binding
domain on enzyme activity
Microbiology 146 (Pt 11), 2947-2955 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Caldibacillus cellulovorans
Caldibacillus cellulovorans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae;
Caldibacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon start=1
/transl_table=11
/product="beta-1,4-xylanase XynA precursor"

    .3237
    .0rganism="Caldibacillus cellulovorans"
/mol_type="qenomic DNA"
/db_xref="taxon:74586"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 (bases 1 to 3237)
Sunna, A., Gibbs, M.D. and Bergquist, P.L.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'note="multidomain protein"
  Gaps:
                                                US-09-917-376-5 (1-88) x AF163837 (1-4567)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        295 GCGGACTACTATCGAG 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <1. .279
/note="orf1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       384. .3149
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20519260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11065373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF200304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
DB:
                                                                                                    ò
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                     IHFTINMALYNETDDWCCNICK, CONTRACTION TENGEL COSSIPPRIVATE PROTECT TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED THE TRANSPINED TO THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED TO THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED THE TRANSPINED T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /translation="MGASGVFSRTAQNGYGLKARCQKRRCVGRNLLGHRLFVVDDDEF
FFAGARKALGRFSBRQNRLGBAQNFBSGTVTFVQPDDKRRIBRBNIGQLFVBFQASI
GPKRLIEJAEQRQARARKQASDQPQLKNAVILHFIDQHMPYGVBRMDFFHKFIBKQKS
RSILTRYY"
                                                                                                                                                                          'note="Region: beta-1,4-mannanase catalytic domain"
'evidence=experimental
                                                                                                    product="multidomain beta-1,4-mannanase precursor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                note="Region: cellulose-binding domain type IIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         note="Region: cellulose-binding domain type IIIb"
evidence=not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             product="multidomain beta-1,4-mannanase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               note="Region: proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17 27 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 evidence_not_experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /evidence=not_experimental
719. .3409
                                                                                                                         /protein_id="AAF22274.1"
/db_xref="GI:6651327"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /protein_id="AAF22275.1"
/db_xref="GI:6651328"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'product = "unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             codon_start=1
transl_table=11
                                                                              table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1349. .1480
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .928. .2032
'gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2033. .2884
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2885. .2959
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2960. .3409
/gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3894. .4400
/note="orf3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'gene="manA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1481. .1927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.3e-17
225.50
67.44%
47.67%
48.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              620. .718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
SOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
```

BCT 15-JUN-2002

ŏ

```
DOYNGAIL ACCOUNT AGAIN TO THE ATT TO THE ACCOUNT AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN AGAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DYNDGVGI VTEBERKDY FRY PAKE I YLPANWSGTFGQGNT I PGTAGVPSDPAKGGNGV
Y I GYSDLRPA I KQDPAWAYLDNKYKTSYNPTTKQWENGAPTFTYHR FWSQVDMATAYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SVATGVTGSTFTDTĞLTNGTAYYYVI TAVNSAGESAPSAQVTATPQAGVTVPGVFTLS
GMAGDAQAVLNWTASTGAATYTVQRSTGTGAYANLATGLTALTYTDTTAVNGTAYNYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EFDRLLGDSSTPGVQVPSAPAGVTAAGGDKQVVLSWNAAAGATSYSVKRATVNGGPYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                //ATNASGQTNSNVVVLTPMAAPISTGTLEVQYRNGGSSASGNAITPQFNLKNTGTTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IDLSKVKVRYYFTKDSAADMSFWCDYAQLGSANVQGSFVAVNPAKGTADTYLELSFSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAGSLAAGAETGI I QTRFSKNNWTNFDQSNDYSFNATQTAFGAWTKVTAYQDGAKVWG
                                                                                                                                                                                                                                    ce148C gene; cellulose 1,4-beta-cellobiosidase.
Paenibacillus sp. BP-23
Paenibacillus sp. BP-23
Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Direct Submission
Submitted (31-MAY-2002) Diaz P., Microbiology, University
Barcelona, Av. Diagonal, 645, 08028-Barcelona, SPAIN
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                  Sanchez,M.M., Pastor,F.I.J. and Diaz,P.
Paenibacillus sp. BP-23 family 48 cellulase. Cloning and
performance on cellulosic substrates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene="ce148C"
167. .3334
/gene="ce148C"
/product="ce11ulose 1,4-beta-ce11obiogidase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /country="Spain:Catalonia, Ebro river delta"
                                                             3509 bp DNA linear
BP-23 cel48C gene for cellulose
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .3509
/organism="Paenibacillus sp. BP-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /EC_number="3.2.1.91"
/function="cellulose hydrolysis"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /mol_type="genomic DNA"
/strain="BP-23"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /db_xref="taxon:198119"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          table=11
                                                                                          Paenibacillus sp. BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /gene="cel48C"
20. .25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /gene="ce148C"
39. .45
                                                                                                                                                                                                          AJ488933.1 GI:21449823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene="cel48C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /gene="cel48C"
/EC_number="3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /gene="cel48C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (bases 1 to 3509)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .3337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /trans]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Diaz, P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -10_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 promoter
                                                                    LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
RESULT 9
PAE488933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gene
                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EATURES
                  / translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Margradge" | translation="Mar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2694 GGCAACCTGGTCGTGCAGTACCGCGGGGGGAACGCGGGGGGAAACCTGGTGAAG 2753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2754 CCGCATTTTCGGATTGGAACGGGACGAGCGTACCGTTGTCGAGCTTACGATT 2813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RVEYRVGDSSATDNOMKPQLRIVNTGSQAVPLTELKVRYWYTKNSTQAEQYFCDWAQI
                                                                                                                                                                                                                                                                                                                                                                                                                     GCSNIRAQFVSLAQPVSGADSYIELSFTGGSVPAGGNTGEIQNRIHFTNWMNYNETDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTTSVPLSELTIRYWYTVDGDKPQVFNCDWAWVGCSNLRGSLVKLTTGRTGADYYLEI
TFTSGAGSLAPGANSGDIQARINKNDWTNYNEANDYSYDPTKTSFADWNRVTLYRNGQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   43 ArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAla 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     _note="Region: beta-1,4-xylanase catalytic domain"
947. .2075
gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cellulose-binding domain typeIIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note="Region: cellulose-binding domain typeIIIb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       thermostabilizing domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                proline-threonine linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3237
40
16
29
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 product="beta-1,4-xylanase XynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-917-376-5 (1-88) x AF200304 (1-3237)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2931 GCGGACTACTACCTTGAG 2948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="Region: c
2520. .2696
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       note="Region:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                note="Region:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note="Region:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     483. .3146
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384. .482
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83. .965
gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    966. .1946
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.99e-16
214.50
65.12%
46.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2076. .2519
/gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene="xynA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2697. .3146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45.74%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CVWGVEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . . No. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

ଚ 셤 ଧ g Š 셤 ò 셤 ò a

```
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                             sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
E35142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                  3056 GCACAGCTTGGCAGTGCCAATGTGCAGGCAGCTTTGTCGCTGTAAATCCGGCGAAGGGC 3115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pastor, F.I., Pujol, X., Blanco, A., Vidal, T., Torres, A.L. and Diaz, P. Molecular cloning and characterization of a multidomain endoglucanase from Paenibacillus sp BP-23: evaluation of its performance in pulp refining Appl. Microbiol. Biotechnol. 55 (1), 61-68 (2001)
                                                                                                                                                                                                                                                2879 ACGGGAACGCTGGAAGTGCAGTATCGCAATGGAGGGTCCAGTGCTTCAGGTAATGCAATT 2938
                                                                                                                                                                                                                                                                                             2939 ACGCCGCAATTTAATCTGAAAAATACAGGTACAACGGCAATTGATCTGAGTAAGGTGAAG 2998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BCT 09-MAY-2001
                                                                                                                                                                                                                                                                              LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                           AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                     SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Direct Submission
Submitted (10-MAR-1999) Diaz P., Microbiology, University of
Barcelona, Av.Diagonal 645, Barcelona-08028, SPAIN
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria, Firmicutes; Bacillales; Bacillaceae; Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                     3509
40
21
25
1
                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AJ133614.1 GI:4490765
celB gene; cellulase; glycosyl hydrolase.
Bacillus sp. BP-23
Bacillus sp. BP-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. .4161
/organism="Bacillus sp.
/mol_type="genomic DNA"
/strain="BP-23"
                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4161 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref="taxon:89769"
398. .3525
/gene="celB"
                                                                                                                                                                                        US-09-917-376-5 (1-88) x PAE488933 (1-3509)
/EC_number="3.2.1.91"
3456. .3472
/gene="cel48C"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3116 ACGGCAGATACGTATCTGGAG 3136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BSP133614 4161
Bacillus sp. BP23 celB gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  398. .404
/gene="celB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                137. .442
/gene="celB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         514. .519
/gene="celB"
532. .3525
                                                                                    9.26e-16
209.50
70.11%
45.98%
44.67%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11234960
2 (bases 1 to 4161)
                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AJ133614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Diaz, P.
                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -35_signal
                  terminator
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -10_signal
                                                                                                                                                                                                                      ~
                                                                                                                                                                                                                                                                              22
                                                                                                                                                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                                                                                                                                                                      82
                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BSP133614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RBS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PEATURES
                                             ORIGIN
```

```
PSYQNLTTGNPVKSNYI PVYDAGVKVSGQEPGYTPVAVPAAPAGVQAVAGNGOVALLW
SASAGAVSYTVKRAEVSGGPYTTVAAGVNGLTYTNTGLTNGKTYYYVVTAVNSAGESP
ASVQOSGNPQDATTVPGAVTLSGTAGNNGNNLSWTAAGSASSYTVQRAVAGGTYTDVA
TGLAVLNVNDTTALNGTSYSYRVAAVNASGQTLSNVVMLTPSGPPVSTGTLEVQYRSG
GSGNSSNAVTPOPNLKNTGTQAIDLSTVKIRYYFTKGGTBELSFWCDYAQVGSANVQG
MPVAVNPAKGTADTYVEISFTSGAGSLAAGAETGVIQTRFSKNNWSAFDQSNDYSYDA
SKTAFAAMNKVTAYQGNTQVWGLEP"
                                                                                                                                                                                 /transl_table=11
/product="cellulase"
/productan_id="CAB18941.1"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.092411"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/db_xref="GOA.09241"
/d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELYNPADTYRCKYTDCITDAAAFYNSWTGYEDELAWGGAWLYLATNDNAYLSKALSAA
DRWSTSGGSANWPYTWTGGWDSKHYGAQILLARITSNLNMPEATKFIQSTERNLDYWT
VGTNGGRVKYTPGGLAWLDQWGSLRYAANAAFISFVYSDWVSDPVKKSRYQNFATSQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NY ILGDNPROSSYVYGYGQNSPQHPHHRTAHSSWMNNEDIPANHRHILYGÄMYGGPNA
SDQYTDDIGVYSNBYATDYNAGFTALAKMILYGQNHQPLANFPABEYKGDBYFVE
AAVRSGSSNYTEIRALLINNRGSWPARMGDQLSFKYFLDLSBYYAAGRTVSDVQYTVSS
SEGATYSQPVVVDAARKIYAITANFSNYKIYPGGEGNYRKEVQPRITGPOGAMNPAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3067 ACAGGTACGCTTGAGGTGCAGTATCGAAGCGGGGGGTTCAGGCAATTCCAGCAATGCAGTT 3126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3127 ACTCCGCAATTCAATCTGAAGAATACCGGCACACAGGCGATTGATCTGAGTACGGTGAAA 3186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3187 ATTCGTTATTACTTTACCAAAGAC---GCACGAGGAGGAGCTGTTCTGGTGTGATTAT 3243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT 18-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4161
38
21
27
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
/gene="celB"
/EC_number="3.2.1.4"
/standard_name="endoglucanase"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2029 bp
Truncated cellulase composition.
E35142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x BSP133614 (1-4161)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3304 ACAGCGGATACGTATGTGGAG 3324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4.83e-15
204.50
67.82%
43.68%
43.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                532. .633
/gene="celB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E35142.1 GI:13018967
JP 1999221086-A/44.
```

```
source
                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 of 5
                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                               SELL
                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF078038S5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SEGMENT
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                     $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ծ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             원
Paiji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G.,
Hyu,M. and Daian,P.W.
Truncated cellulase composition
Datent: JP 199221086-A 44 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
OS Artificial Sequence
N JP 1999221086-A/44
PD 17-AUG-1999
PF 21-SEP-1999 UP 1998283606
PR 19-SEP-1997 US 08/932571
PI PAIJI ANDERSON, DETAER L BAGUKUISUTO,ROY M DANIEL, PI
GURAHAMY FARINTON,
PI MORELAND DAVID GIBUSU,HYU MORGAN,DAIAN PURATONIOTISU WILLIAM
PC CIRNIS/09,CI1D3/386.CI2N1/21,CI2N9/42//(CI2N1/21,CI2R1:19), PC
(CI2N9/42,CI2R1:19),CI2N15/00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT 18-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||:::|||
| ATAAGGCCGTGGTTTAAGATAGTGAATGGAGGCAGCAGCAGTGTTGATCTTAGCAGGGTT 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 AAGATAAGATACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||||||| :::|||
178 TGGGCACAGATAGGGGCAAGCAATGTGACATTTTGTGAAGCTTAGCAGGGGGGTG 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 (bases 1 to 2029)
Paiji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G.,
Hyu,M. and Daian,P.W.
                                                                                                                                                                                                                                                 . .2029 organism='Artificial Sequence'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                         2029
38
15
15
1
                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                         1. 2029
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="teaxon:32630"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Truncated cellulase composition
Patent: JP 1999221086-A 45 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
OS Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2029 bp
Truncated cellulase composition.
E35143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               238 AGTGGAĞCĞĞATTATTACCTĞGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x E35142 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               E35143.1 GI:13018968
                                                                                                                                                                                                                                                                                                                                                                                        2.6e-14
195.50
60.23$
43.18$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JP 1999221086-A/45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unidentified
unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unclassified
                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                    임료답답
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                           . No. .
                                TITLE
JOURNAL
     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                             FEATURES
                                                                          COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMMENT
                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E35143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
PN JP 1999221086-A/45
PD 17-AUG-1999
PF 21-SEP-1999 JP 19982813606
PR 19-SEP-1997 US 08/932571
GURAHAMU K FARINTON,
PI MORELAND DAVID GIBUSU, HYU MORGAN, DAIAN PURATONIOTISU WILLIAM
PC C12N15/09, C11D3/386, C12N1/21, C12N9/42// (C12N1/21, C12R1:19), PC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6005 bp DNA linear BCT 11-FEB-2001 Caldicellulosiruptor sp. Tok/B.1 glycosyl hydrolase 6 gene, partial cds; and CelE gene, complete cds.
AF078042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 ATAAGGCCGTGGTTTAAGATAGTGAATGGAGGCAGCAGCAGCAGTGTTGATCTTAGCAGGGTT 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 AAGATAAGATACTGGTACACAGTGGTGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||||||| :::|||
178 TGGGCACAGATAGGGGAAGCAATGTGACATTTTGTGAAGCTTAGCAGCGGAGTG 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspAsrAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ATGGGAAGTGGTGTGAGGTACTGTACAAGAGAAGATGAGACAAGTGCGAGCACAGGTTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          from the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 6005)
Gibbs,W.D., Reeves, R.A., Farrington, G.K., Anderson, P., Williams, D.P. and Bergquist, P.L.
Miltidomain and multifunctional glycosyl hydrolases from extreme thermophile Caldicellulosiruptor isolate Tok7B.1 curr. Microbiol. 40 (5), 333-340 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10706665
2 (bases 1 to 6005)
Gibbs,M.D., Reeves,R.A., Farrington,G.K., Anderson,P.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Caldicellulosiruptor sp. Tok7B.1
Caldicellulosiruptor sp. Tok7B.1
Bacteria; Firmicutes; Clostridia; Clostridiales;
Syntrophomonadaceae; Caldicellulosiruptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2029
38
15
34
1
                                                                                                                                                                                                                    Location/Qualifiers
1. .2029
/organism='Unidentified'
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                 1. .2029
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    238 AGTGGAGCGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81 ProThralaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-5 (1-88) x E35143 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2.6e-14
195.50
60.23%
43.18%
41.68%
```

```
1. .6416
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  fruncated cellulase composition.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-5 (1-88) x AF078038S5 (1-6005)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                   3107. .3223
/note="PT-linker"
4196. .4295
/note="PT-linker"
   /note="PT-linker"
3107. .3223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E35100.1 GI:13018925
JP 1999221086-A/2.
unidentified
unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2885 GATTATTACCTGGAG 2899
                                                                                                                                                                                                                               1.36e-13
194.50
61.18%
44.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                               misc_feature
                                                                                                   misc_feature
                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84
                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E35100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLITION TO THE LAND WITH THE TOTAL THE THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TOTAL THE TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EPGRNAQI KWWYANGNIGSMTWULNPKI KI ENVGTTAVDI.SRVKNRYWTI DGEATQS
YSTOSI NPAY I DVKPWI.GANAGADYYBE GREGAGALAAGOSTREI ILGI ÇIĞG
SYNGSINDYS YRBATI YI ENEKUTGYI DGA I WAREPESRGTKBAĞOYTPT PAPTPTST
PTPTPTTPTPTPTTY I ENEKUTGYI DGA I WAREPESRGTKBAĞOYTPTPAPTPTST
PTPTTPTPTPTTTPTPTY VYTPTPAYTPDVKI SI DTSRGRTKI SPYI YGANQDI QGVVH
PTPTPTTPTPTTPTTY YORDANGANGON YHISSDDYWYT IMGITTADIK NYDAAVUS KFRED
SI KONAY SA I TLQMUGYVA KOĞNGTVSESETAR PŞRARAVK FKDĞALSI, QPDYNDINY
VYMDEFINYLI NKYĞRS SSATĞI KĞYI LLDNEPDLMFTTHPR I HPQKYTCSELI NKSYB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LAKVIKTLDPDAEIFGPASYGFVGYLTLQDAPDWNQVKGNHRWFLSWYLBQMKKASDS
FGKRLLDVLDIHWYPBAQVGGVRICFDGENSTSRDVAIARMQAPRTLWDPTYKTTQKG
QITAGENSWINQWFPEYLPLLPNIKADIDKYYPGTKLAITEFDYGGKDHISGGIALAD
VLGIFGKYGVYMAARWGDSGSYAQAAYNIYLNYDGKGSRYGSTCVSAETTDVENMPVY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DLFEFADTTRSDAGYTAATGFYTSGGFIDDLGWAAVWLYIATNDSSYLTKAEELMSEY
ANGTNTWTQCWDDVRYGTLIMLAKITGKELYKGAVERNLDHWTDRITYTPKGMAYLTG
WGSLRYATTAAFLACVYADWSGCDSNKKTKYLNFAKSQIDYALGSTGRSFVVGFGTNY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein_id="AAK06394.1"
db_xref="GI:12743885"
'translation="MQEMKAIKRVVSITALLVLTLSLCFPGIMPVKAYAGGTYNYGEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOKT I MFY BFOMSGKL PSWVRNNWRGDSGLDDGKDVGLDL TGGWHDAGDHVK FNLPMS
YSASMLGWAVY EYKDA FVKSKOLEH I LNO I EWANDY FVKCHPSKYVY YYQVGDPTVDH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NFWGPAEVMOMKRPAYKCDLSNPASSVVAETAASLAVASVVIKERNSOKAASYLQHAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASIEGEDDSTVHIILINRNYDRKLKAEIKMNNTRVYTGGEIYGFDSTSSQIRKMGVLS
Williams, D.P. and Bergquist, P.L.
Direct Submission
Submitted (27-OCT-1999) Biological Sciences, Macquarie University,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1224. .4195
note="duplicated type IIIb cellulose binding site"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         <1. .503
/note="Region: family 48 glycosyl hydrolase domain"</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25. .2053
'note="Region: family 9 glycosyl hydrolase domain"
                                                                                                                                                                                         organism="Caldicellulosiruptor sp. Tok7B.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2054. .2545
/note="type IIIc cellulose binding site"
2627. .3106
/note="type IIb cellulose binding site"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         117. .5872
'note="multidomain glycosyl hydrolase"
                                                                                                                                                                                                                                                                                                                                                                                                         transī table=11
product="glycosyl hydrolase 6"
protein_id="AAK06393.1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NIQNNTITIEVPNLTVYHIVLTSSK"
                                                                                                                                                                                                                                                                                                                                            note="ORF6; multidomain"
codon_start=3
                                                                                                                                                                                                                      /mol_type="genomic DNA"
/isolate="Tok7B.1"
                                                                                                                                                                                                                                                                               db_xref="taxon:80339"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   evidence=experimental
                                                                                      Sydney, NSW 2109, Australia
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               transl table=11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               product="CelE"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         codon_start=1
                                                                                                                                                   .6005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EILFGNO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                   TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                            FEATURES
```

```
MORELAND DAVID GIBUSU, HYU MORGAN, DAIAN PURATONIOTISU WILLIAM C12N15/09, C11D3/386, C12N1/21, C12N9/42//(C12N1/21, C12R1:19), PC (C12N9/42, C12R1:19), C12N15/00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2648 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAGGCCG 2707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2708 TGGTTTAAGATAGTGAATGGAGGAGCAGCAGTGTTGATCTTAGCAGGGTTAAGATAAGA 2767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |||||||:::|||
2768 TACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT 18-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 (bases 1 to 6416)
Paiji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G.,
Hyu,M. and Daian,P.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OS UNKNOWN

PN JP 1999221086-A/2

PD 17-AUG-1999

PF 21-SEP-1999 UP 1998283606

PR 19-SEP-1997 US 08/932571

PI PAIJI ANDERSON, PETAER L BAGUKUISUTO, ROY M DANIEL, PI GURAHAMU K PARINTON,

PI MORELAND DAVID GIPTING TIPLICATOR TO THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
6005
38
14
32
1
      Length:
Matches:
Conservative:
Mismatches:
Indels:

    6416
    organism='Unknown'.

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Truncated cellulase composition
Patent: JP 1999221086-A 2 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
```

```
BCT 22-OCT-1997
complete cds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FIYNLIKSLKEKGVPIHGVGLQCHINLDWPSISEIENTIKLFSSIPGIEIHITELDMS
SYGOSSTYNTAVPPRELLIKQAQRYEKELPDLRKKKNVVTNTFWGLKODYSWLSKNY
GKCOYPLLEDENYMSKYAFWSLIDSSIVPTTTLPTPPEDLQLPTTFPPTASVTPTPT
VTPTPTVTVTPTPPTPTPTTATPTPPTPTPTSPASGQIKKLYANKETNSTTNTIR
                                                                                                                                                                                   SINIMT I RGHTLVWHQQTPDWFPKDSNGNTLSKDALLSRLKQY I YDVVGRYKGKVYAWD
VVNEA I DEGQSDGYRRSNWYN I CGPEY I EKAFI WAHEADPNAKL FYNDYNTENSQKRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    translation="MEKKLILKHFNSITPENEIEPSELLINENTYNFSKADSFVNFAI
                                                                                                                                                                     23
                                                                                                                                                                                                                                     43
                                                                                                                                                                                                                                                                                                  TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                       83
                                                                                                                                                                GlyvalLysvalGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                  GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                   64 MetglyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Miyake,K., Machida,Y., Hattori,K. and Iijima,S.
Characterization of a multi-domain cellulase from an extremely
thermophilic anaerobe strain NA10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Leases 1 to 4743)
Miyake,K., Machida,Y. and Iijima,S.
Direct Submission
Submitted (09-007-1997) Katsuhide Miyake, Nagoya University,
Department of Biotechnology, School of Engineering,; Furo-cho,
Chikusa-ku, Nagoya, Aichi 464-01, Japan
(E-mail:miyake@proc.nubio.nagoya-u.ac.jp, Tel:+81-052-789-4278,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABOUB029 4743 bp DNA linear
Thermophilic anaerobe NA10 gene for beta-glucanase,
AB008029
                 6416
38
14
32
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="thermophilic anaerobe NA10"
/mol_type="genomic DNA"
/strain="NA10"
             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /codon_start=1
/trans1_table=11
/product="beta-glucanase"
/protein_id="BAA22939_1"
/db_xref="G1:2564015"
                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /db_xref="taxon:67756"
336. .3338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                               US-09-917-376-5 (1-88) x E35100 (1-6416)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          thermophilic anaerobe NA10 thermophilic anaerobe NA10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2902 GATTATTACCTGGAG 2916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AB008029.1 GI:2564014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon Ft**
/tr*
                                                                                                                                                                                                                                                                                                                                                                                                                                       84 AspThrTyrLeuGln 88
               1.47e-13
                             194.50
61.18%
44.71%
41.47%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 . .4743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           beta-glucanase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unpublished
                                                            Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (sites)
                                                Percent Similarity:
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                  24
                                                                                                                                                                                                                                                                                                     44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
AB008029
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                ò
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
GVWSCNIKSTLAEIANRGFNLLRVPISAELILNWSQGIYPKPNINYYVNPELEGKNSL
EVFDIVVQTCKEVGLKIMLDIHSIKTDRAMGHYFVEWYDEKFFPEDFYKACEWITNRYK
NDDTIIAFDLKMEPHGKFWQDTTPAKWDNSTDINNWKYABETCAKRILNINRYLLIVI
EGIEAYPKDDVTWTSKSSSDYYSTWWGGNLRGVRKYPINLGKYQNKVYSPHDYGPSV
YQQPWFYPGFTYESLLQDCWRPWAYIMEENIAPLLIGEWGGHLDGADNEKWYKZLRD
YIIENHIHHTFWCENANSGDYGGLYGYDFTTWDEKKYSFLKPALWQDSQGRFVGLDH
RPLGINGKNINITTYYNNNEDEPVPATK"
PWLKVVNTGSSSIDLSRVTIRYWYTVDGDKAQSAVSDWAQIGASNVTFKFVKLSSSVS
BADYYLLEIFEKSGAGQLQPGKOTGEIOTRFNKEDWSNYNQGONDWSWIGSMTSVERDMK
VTAYIDGVLVWGQEPTGATAAPIATPFTPTPAPTSTPTPTPAAPITTPFTTT
TATPAPTPAPTSTPAYLDDTNDDWLYVSGNKIVDKDGRPVWLTGVNWFGYNTGTNVFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1626 GCACAGATAGGAGCAAGCAATGTAACATTCAAGTTTGTGAAGCTGAGCAGTAGCGTAAGT 1685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1449 AGCGGACAGATAAAGGTACTGTATGCTAACAAGGAGACAACAGCACGACAAGACACGATA 1508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x AB008029 (1-4743)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   completed: May 12, 2004, 09:26:47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1686 GGAĞCĞĞACTATTACTTĞGAG 1706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                1.79e-13
192.50
63.22%
43.68%
                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Job time : 1298.92 secs
                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search
                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . 음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
```

Bacillus NK-1 cell Endol gen

CMCase

Pectate 1 C. thermo

Nucleotid Nucleotid

Bacillus

Prokaryot

Fusion pr Endo3 gen

C. cellul Clostridi Cellulose

Clostridi

Fusion pr Fusion pr

DNA encod DNA encod

Perfect score:

Sequence:

ı

OM protein

Run on:

Scoring table:

Total number

Searched:

```
mannanase A; glycoside hydrolase; enzyme; hemicellulose; sugar;
feed; paper pulp; biofuel; mannase; gene; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Aad11044 (Aad11045 (Aad11043 (Aaz27550 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Aca25846 | Aaz92616 (Ada02720 |
                                        Aax55661
Aad2625
Aad2625
Abx52202
Abx52202
Aag49820
Aag49820
Aag49820
Aag49820
Aag49820
Aag49820
Aag49820
Aag49820
Aag49820
Aag49300
Aag4747
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac27475
Aac274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Aad22201
Aaf12881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nucleotide sequence of the ManA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                             AAD26525
AAL41028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAF86248
ABK52403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAD11046
AAD11044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1. .2289
/*tag= a
/product= "ManA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX24930
                                                                                                                                                                                                                                                                                                                                                                                                                                           AAQ13003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAD11042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACA25846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADA02720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ΤB,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABZ77633 standard; DNA; 2289 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-2001; 2001WO-US023819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vinzant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
(MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ws,
WO2003012110-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 03-JUN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-FEB-2003
194.5
194.5
192.5
192.5
188
188
188
186
186
165
165
165
165
165
                                                                                                                                                                                                                                                                                                         155.5
150.5
130.5
126.5
126.5
126.5
120.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABZ77633;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         food;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ManA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABZ77633
  Command line parameters:
-MODEL=frame+ pln.model -DEV=xlh
-MODEL=frame+ pln.model -DEV=xlh
-MODEL=frame+ pln.model -DEV=xlh
-D=/CODE_1/USPTO spool/USO9917376/runat_11052004_114534_28231/app_query.fasta_1.1429
-D=/CODEXT=0 -UNITYS=D104 -QFWT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITYS=D108 -QFWT=1 -END=-1 -MATRIX=D10sum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=100 -THR SCORE=pct -THR NAYE-100 -THR MING=15
-MODBE_LOCAL -OUTFWT=pto -NORM=ext -HEAPPSIZE=500 -MINLENE O -MAXLEN=2000000000
-USER=US09917376_@CGN 11_1758_@runat_11052004_114534_28231 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NGG SCORES=0 -WALT -DSPBELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -FGAPOP=6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nucleotid
Nucleotid
Acidother
A. cellul
Portion o
DNA seque
Active ce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleotid
                                                                                                                                                 ; Search time 122.068 Seconds (without alignments) 3062.577 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                            469
1 VSGGVKVQYKNNDSAPGDNQ......IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Abz77634 Pbz77632 Padd22922 Padd22922 Padd2292 Padd15178 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568 Padd26568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Abz77633
                    GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                        nucleic search, using frame plus p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3373863 segs, 2124099041 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABZ77633
ABZ77634
ABZ77632
ADD22922
ABZ76162
AAQ15178
AAXS5660
                                                                                                                                                      May 12, 2004, 03:07:10
                                                                                                                                                                                                                                                                                                                                                                           , Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                                                                                       Xgapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         geneseqn2003bs:*
geneseqn2003cs:*
geneseqn2004s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          geneseqn2001bs:*
geneseqn2002s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       geneseqn2001as:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       geneseqn2003as:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   geneseqn1980s:*
geneseqn1990s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N_Geneseq_29Jan04:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   geneseqn2000s:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                       US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match Length DB
                                                                                                                                                                                                                                                                                                                           BLOSUM62
Xgapop 10.0 , x
Ygapop 10.0 , y
Fgapop 6.0 , F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3687
2869
2869
3365
2600
2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0
```

5: 7:: 9:: 10:

Jatabase

466 462 462 436 231.5 195.5

2645978

Score

Result Š.

DNA seque E. coli c Chimeric

Clostridi

Clostridi Clostridi

Gaussia Gaussia Prokaryot Genomic D Human NFA

Aspergill

Decker

##X####X

```
1930 TGGGCGGCGATGGGCGTGTGGGAATATCCGCCCTCGTTCGGCTCGGTGAACCCGGCGACG 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGCGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1810 ATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGTGGATTTGTCGACGGTG 1869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1870 ACGGGCGTACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC 1929
                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present sequence encodes a GuxA polypeptide. GuxA is thermostable cellulase, and is a member of the glycoside hydrolase family of enzymes. GuxA is useful for reducing cellulose in a starting material such as agricultural biomass to sugars. This is useful in biofuel production. GuxA is also useful in the conversion of biomass to biofuels and biofuel additives, in detergents, pulp and paper processing, food and feed processing, and in textile process. GuxA is also useful for raising polyclonal and monoclonal antibodies that are useful in purifying GuxA, or detecting GuxA polypeptide expression, as well as reagent tools for
                                                                                                                                                                                                                                                                                                                                 Novel thermal tolerant GuxA polypeptide derived from Acidothermus cellulolyticus, useful for reducing cellulose in a starting material, for the conversion of biomass to biofuels and biofuel additives.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 3687 BP; 725 A; 1173 C; 1171 G; 618 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 characterizing the molecular actions of GuxA polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                 Himmel ME,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-5 (1-88) x ABZ77634 (1-3687)
                                                                                                                                                                                                                                                                                                                                                                                                                    Example 1; Page 23-24; 47pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProThrAlaAspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CCGACGGCGACACCTACCTGCAG
                                                                                                                                                                                                                               Vinzant TB,
                                                                                                        28-JUL-2001; 2001WO-US023817.
                                                                                                                                                28-JUL-2001; 2001WO-US023817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABZ77632 standard; DNA; 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.79e-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  466.00
100.00%
98.86%
99.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-JUN-2003 (first entry)
                                                                                                                                                                                                                                                                       2003-239526/23
                                                                                                                                                                                         (MIDE ) MIDWEST RES
                                                                                                                                                                                                                               Ding S, Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity:
                                                                                                                                                                                                                                                                                             P-PSDB; ABP73029
                         WO2003012109-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                               13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABZ77632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABZ77632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RXXXXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1483 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGTCGTCGACACTGGTGTACAACTGTGAAC 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present sequence encodes ManA, a thermostable mannanase A polypeptide derived from Acidothermus cellulolyticus. ManA is a member of the glycoside hydrolase family of enzymes. ManA is useful for reducing hemicellulose in a starting material to simpler carbohydrate units, and ultimately to sugars which are useful in the food, feed, paper pulp, and biofuels industries. It is useful for the processing of food and in food stuffs as bulking agents, and for the degradation of mannase. ManA is also useful to raise polyclonal and monoclonal antibodies that are useful in purifying ManA, or detecting ManA polypeptide expression, and as well as reagent tools for characterizing the molecular actions of ManA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GuxA; cellulase; glycoside hydrolase; enzyme; cellulose; sugar; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                           Novel thermal tolerant mannanase A polypeptide derived from Acidothermus cellulolyticus, useful for reducing hemicellulose in a starting material, for processing of food, and as bulking agents in food stuffs.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValSerGlyGlyValLysValGInTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1363 GTGTGGGGTGGGGTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Arcaaaccagercrearregranaracagarcarcarcarcarrarrarcaacara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                   BP; 463 A; 700 C; 688 G; 438 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2289
0 0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nucleotide seguence of the GuxA polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x ABZ77633 (1-2289)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1. .3687
/*tag= a
                                                                                                                                              Example 1; Page 23; 46pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product= "GuxA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABZ77634 standard; DNA; 3687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4.13e-48
469.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
WPI; 2003-248182/24.
P-PSDB; ABP73022.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2289
                                                                                                                                                                                                                                                                                                                                                                                                 polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABZ77634;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
```

ò g ò g ò g ò 셤 ò

40

9

80

Key

RESULT 2 ABZ77634

요

m

```
2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                       61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                              Acidothermus cellulolyticus avicelase AviIII DNA
                                                                                                                                                                                                                                                  glycoside hydrolase; avicelase; ds; gene
                                                                    ADD22922 standard; DNA; 2869 BP
                                                                                                                                                                                                                                                                                                                                                                                                       18-OCT-2002; 2002US-00155400
                                                                                                                                                                                                                                                                        Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                          US2003108988-A1
                                                                                                                                                                                        15-JAN-2004
                                                                                                                                                                                                                                                                                                                                                                                 12-JUN-2003
                                                                                                                                                                 ADD22922;
                                                                                                                    RESULT
                                                                                                                                                       셤
                                                                     ઠે
                                                                                           셤
                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence encodes a thermostable avicelase polypeptide, designated AviIII. AviIII is a member of the glycoside hydrolase family of enzymes, and is a cellulase. AviIII is meatul in the conversion of biomass to biofuels and biofuel additives. It may be useful in the processing and in textile processes. The thermostable AviIII peptide is useful in the degradation of cellulose, and in generating specific anti-polypeptides that are useful in purifying recombinant AviIII polypeptide from genetically engineered host cells, in detecting AviIII polypeptide expression, as well as a reagent tool for characterizing the molecular actions of the polypeptide. The AviIII polymucleotide is useful
                                                                                                                                                                                                                                                                                                                                                                   New thermostable AviIII peptide from Acidothermus cellulolyticus, useful for degradation of cellulose or in generating anti-AviIII antibodies for purifying recombinant AviIII polypeptides from genetically engineered host cells.
                     Avicelase; AviIII; glycoside hydrolase; enzyme; cellulase; biofuel; detergent; pulp processing; paper processing; feed processing; textile;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   as a source of probes or primers in various diagnostic assays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                      /transl_except= (pos:2869,aa:Xaa)
/note= "Xaa is an unspecified residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2869
86
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
 sequence of the avicelase AviIII.
                                                                                                                                                                                                                                                                                                           Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-5 (1-88) x ABZ77632 (1-2869)
                                                                                           Location/Qualifiers
1. .2869
/*tag= a
/partial
/product= "AviIII"
                                                                                                                                                                                                                                                                                                           Vinzant TB,
                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 3; Page 24; 44pp; English.
                                                                                                                                                                                                                                                             28-JUL-2001; 2001WO-US023818
                                                                                                                                                                                                                                      28-JUL-2001; 2001WO-US023818
                                                                    Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.09e-47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      462.00
100.00%
97.73%
98.51%
                                                                                                                                                                                                                                                                                   (MIDE ) MIDWEST RES INST
                                   detergent; pulp proce cellulose; gene; ss.
                                                                                                                                                                                                                                                                                                                                 WPI; 2003-248177/24.
                                                                                                                                                                                                                                                                                                           Adney WS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity:
                                                                                                                                                                                                                                                                                                                                               P-PSDB; ABP73015.
                                                                                                                                                                                        WO2003012090-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                               13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Н
                                                                                                                                                                                                                                                                                                           o,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
80
..
                                                                                                                                                                                                                                                                                                           Ding
                                                                                             Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
 ò
```

cellulose reduction; agricultural biomass; municipal solid waste;

(first entry)

Location/Qualifiers 1. 2868 /*tag≈ a /product= "AviIII"

```
New isolated thermal tolerant avicelase polynucleotide useful for detection of a polynucleotide encoding AviIII and for reducing cellulose in a starting material, e.g. municipal solid waste.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The invention relates to an isolated polynucleotide molecule encoding thermostable Avilli polypeptide. The polynucleotide is useful for detection of a polynucleotide encoding Avilli. The polynucelotide is useful for setting to reducing cellulose in a starting material which involves administering to the starting material, e.g. agricultural biomass or municipal solid waste, a polypeptide molecule of the polynucleotide. The method further comprises administering a second polypeptide molecule chosen from the glycoside hydrolase family of proteins. The present sequence represents DNA encoding the Acidothermus cellulolyticus
                                                                                                       ķ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2869 BP; 546 A; 857 C; 900 G; 565 T; 0 U; 1 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                              Himmel ME;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 17; SEQ ID NO 2; 29pp; English.
                                                                                                                                                                                              Vinzant TB,
28-JUL-2001; 2001US-00917376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           462.00
100.00%
97.73%
98.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .09e-47
                                                                                                                                                                                                                                                  WPI; 2003-810853/76.
P-PSDB; ADD22921.
                                                                                                          (VINZ/) VINZANT T B (HIMM/) HIMMEL M E.
                                                                                                                                                                                           Ding S, Adney WS,
                                                   DING S.
ADNEY W S.
VINZANT T E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  avicelase AviIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                   (DING/)
```

2724

9

ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp

41

40

20

2605 GTGTCGGGTGGAGGTGCAGTATAAGAATAATGATTCGGCGCCGGGTGATAATCAG 2664

유 ò 셤

21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTCGTCGATTGTCCACGATG

ò 셤 ò 셤 ò 셤 ò g ò 셤

```
polypeptide is useful in the degradation of cellulose into biofuel, or for conversion of biomass to biofuel additives. It is used in detergents, pulp and paper processing, food and feed processing, and in textile processing. It can also be used alone or in combination with other cellulase or glycoside hydrolases. The novel polypeptide generates alternative cellulase enzymes capable of assisting in the commercial. Scale processing of cellulose to sugar for use in biofuel production. The present sequence represents a A. cellulolyticus Guxl cellulase encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 crccagricgigaaraccegercercercerearricrceaceereaceerecerac 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     232 TGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGACTGGGCGGCGATG 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              292 GGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACGCCGACGCGGGGGAAC351
                                                                                                                                                                                                                                                                                                                                                                                                                                                   TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                           5 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Detergents; pharmaceuticals; deinking; carboxymethylcellulose; ss.
                                                                                                                                                                                             Sequence 3365 BP; 669 A; 1117 C; 1041 G; 538 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Portion of pAEC-1 contg. cellulase AE-1 gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /*tag= a
/product= "cellulase AE-1"
                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                       (1-3365)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
192. .2540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BP.
                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-5 (1-88) x ABZ76162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAQ15178 standard; DNA; 2600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  90JP-00045465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90JP-00045465.
                                                                                                                                                                                                                                               8.56e-44
436.00
97.62%
96.43%
92.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (revised)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCTACCTGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (revised)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (OJIP ) OJI PAPER CO
                                                                                                                                                                                                                                                                                                            Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP03251174-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-FEB-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-AUG-2003
25-MAR-2003
                                                                                                                                                                                                                                 Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-NOV-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17-MAR-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Aeromonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAQ15178;
                                                                                                                                                                                                                                                         No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 6
AAQ15178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Key
                                                                                                                                                                                                                                                                       Score:
   888888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                        TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
|||||||||||||
TGGGCGGCGATCGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGAAC 2844
                                                                                                              2605 Grercegeregasegregaserataagaaraargarregeeeeeergaraareag 2664
                                                                                                                                                                                           2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTGGATTTGTCGACGGTG 2724
                                                                                                                                                                                                                                                                     2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                           ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                        40
                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The invention relates to a thermal tolerant Guxl peptide from A. cellulolyticus. The Guxl exoglucanase is a member of the glycoside hydrolase family and comprises a catalytic domain GH48, carbohydrate binding domain type III, and a carbohydrate binding domain type II. The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New thermal tolerant Guxl peptide having specified amino acid sequence, useful in the degradation of cellulose to biofuels.
                                                                                                                                                                                                                               ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                    | IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         oux.; tnermal tolerant; exoglucanase; glycoside hydrolase; cellulose;
biofuel; detergent; pulp; paper processing; feed processing; textile;
cellulase; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "this codon has an apparent one nucleotide basepair deletion which alters the reading frame"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /transl_except= (pos: 682. .683, aa: Pro)
/note= "this codon has an apparent one nu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Decker SR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WS, Ding S, Vinzant TB, Himmel ME, Mccarter S;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A. cellulolyticus Gux1 protein encoding DNA
     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                       CCGACGGCGGACACCTACCTGCAG 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 22-23; 44pp; English.
                                     US-09-917-376-5 (1-88) x ADD22922 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                     ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. 3365
/*tag= a
/product= "Gux1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    standard; DNA; 3365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-2001; 2001WO-US023820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUL-2001; 2001WO-US023820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Acidothermus cellulolyticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MIDE ) MIDWEST RES INST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2003-300494/29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P-PSDB; ABP71656.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO2003012095-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29-MAY-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                         2845
                                                                             Н
                                                                                                                                                      21
                                                                                                                                                                                                                                                                                                                                                2785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AB276162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABZ76162;
                                                                                                                                                                                                                                 41
                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Adney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lantz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABZ76162
                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
```

24

44

64

```
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAD26568
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                     2202 CTGCGATATTATTTCACAGCTGAC---GGCAATGACCAACTGCAATACAATTGTGACTGG 2258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2259 GCCATGGTAGGCTGCTCCAACCTGAACGGGCTTTCGTGAAAATGAATCCGGGCAAAGCG 2318
                                                                                                                                                                                                                                                                                                                             |||||||
AGCGGGGATTTGGCCGTTCAGTACAAAACCGGAGATACGAATGCAGCCGACAACCAGTTC 2141
                                                                                                                                                                                                                                                                                                                                                                     2142 AAGCCTCATTTCAATATCGTGAACAAGGGGGCTGCGGCAGTGCCTTTAAGCGAGCTGAGT 2201
                                                                                       The plasmid, pAEC 1, was prepd. ligating chromosomal DNA contg. the gene (obtd. from Aeromonas) into pUC18. The gene can be used to produce recombinant enzyme which is used for the effective utilis- ation of biomass resources and the mfr. of pharmaccuticals and foodstuffs, and also for the detergent and delinking of waste paper. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
                                                                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                    81
                                   opt.
                                                                                                                                                                                                                                                                                                              SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                           LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                        ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing; ss.
                                Cellulase AE-1 for e.g. mfr. of pharmaceuticals and foodstuffs - of pH when carboxy:methyl:cellulose is used as substrate.
                                                                                                                                                                        Sequence 2600 BP; 687 A; 649 C; 700 G; 564 T; 0 U; 0 Other;
                                                                                                                                                                                                                      DNA sequence encoding truncated cellulase Cel E3/B5
                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x AAQ15178 (1-2600)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2319 AACGCCGATACCTACTTGGAG 2339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                  Claim 3; Fig 2; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98EP-00810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX55660 standard; DNA; 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-00932571
                                                                                                                                                                                                          1.62e-18
                                                                                                                                                                                                                  231.50
66.67%
48.28%
49.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                  Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-SEP-1998;
                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                        42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                     2082
                                                                                                                                                                                                                                                                                                                                                            22
                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX55660
                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                           ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAX55660
셤
                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                            원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel B1/2, Cel 1/2/3, Cel 6 or Cel B3/B5, or a stability region from one of the defined full-cel or Cel 6 or Cel B3/B5, or a stability region from one of the defined full-cel or K1425 or K1425 or N1426, and Cel B4/5 extends from amino acid K655 to N1426 in the sequence shown in AAV13492; Cel E1 extends from amino acid X93 to D481, Cel E6 extends from x19 to G635, Cel E1/2/3 and the stability region extends from amino acid V133 to K1751 and the stability region extends from amino acid V133 to K1751 and the stability region extends from amino acid V133 to K1751 and the stability region extends from amino acid V139 to G635 in the carymes are useful in laundry detergent compositions to prevent or remove staining backstaining or graying, for use on cellulosic materials including cotton-containing fabries. They are especially useful for preventing redeposition of colorant during stonewashing, and for truncated enzymes show reduced redeposition of dye compared to using nontrinoated enzymes show reduced redeposition of dye compared to using nontrinoated enzymes show reduced redeposition of dye compared to using non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Active cellulase protein; alkalinophilic; textile processing; proteinase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60
                                                                                                                                                             New truncated cellulase proteins, useful in detergents and for producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
Ğ,
  Gibbs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2029 BP; 650 A; 329 C; 520 G; 530 T; 0 U; 0 Other;
Farrington GK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2029
38
15
14
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Active cellulase hybrid protein, E3/B5 DNA.
  Daniels RM,
                                                                                                                                                                                                                                             Disclosure; Page 41-42; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-5 (1-88) x AAX55660 (1-2029)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238 AGTGGAGGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAD26568 standard; DNA; 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3.53e-14
195.50
60.23%
43.18%
41.68%
  ΡΓ,
  , Bergguist F
Williams DP;
                                                                                  WPI; 1999-315403/27.
                                                                                                                                                                                         stonewashed' denim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
                                                                                                           P-PSDB; AAY13494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
  Anderson P,
Morgan H, W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAD26568;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3X8X5X8X8
3X8X5X8X8
```

238 AGTGGAĞCĞĞATTATTACCTĞGAG 261

88

81 ProThrAlaAspThrTyrLeuGln

```
The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the CelBS amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing whits or for imparting stonewashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present
                                                                                                                                                                                                                                                                                                                                                                                                                                        or commercial
detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; B3/B5 hybrid protein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                 New cellulase active protein, useful in textile processing or comdetergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and elevated temperatures.
                                                                                                                                                                                                                                                                                                                                      Θ̈
                                                                                                                                                                                                                                                                                                                                      Gipps
                                                                                                                                                                                                                                                                                                                                      Daniels R,
                                                                                                                 /*tag= a
/product= "E3/B5 hybrid protein"
                                                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                    Bergquist
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Col 71-74; 61pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sequence is E3/B5 hybrid protein DNA
                                                                                 ocation/Qualifiers
                                                                                                                                                                                                                                                                                                  (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                   98US-00136574.
                                                                                                                                                                                                                                                                 97US-00932571.
                                                                                                                                                                                                                                                                                                                                 Anderson P,
                                                                                              .1851
                                                                                                                                                                                                                                                                                                                                              Morgan H, Williams DP
                                                                                                                                                                                                                                                                                                                                                                                   2002-081780/11.
                                                                                                                                                                                                                                                                                                                                                                                                    P-PSDB; AAE16325
                                                                                                                                                                                                                                                                                                                                      Farrington GK,
                                                                                                                                                                                                                                   19-AUG-1998;
                                                  Unidentified
                                                                                                                                                                    US6294366-B1
                                                                                                                                                                                                   25-SEP-2001
                                                                                  Key
```

121 AAGATAAGATACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGAC 177 ||||||| :::||| 178 TGGCACAGATAGGGGCAAGCAATGTGACATTCAATTTTGTGAAGCTTAGCAGCGGAGTG 237 41 ThrvalArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60 1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20 40 9 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80 ::: 1 ATGGGAAGTGGTGAAGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCT IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal Sequence 2029 BP; 650 A; 329 C; 520 G; 530 T; 0 U; 0 Other; 2029 38 15 15 Conservative: Mismatches: Indels: Length: Matches: Gaps: US-09-917-376-5 (1-88) x AAD26568 (1-2029) .53e-14 195.50 60.23% 43.18% 41.68% Local Similarity: Percent Similarity: Alignment Scores: 21 61 Query Match: Score Best ò 임 ò 셤 ò g ઠે g

Conservative: Mismatches: Indels:

Best Local Similarity: Query Match:

Percent Similarity:

US-09-917-376-5 (1-88) x AAX55662 (1-6415)

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B3, Cel B4/5, Cel B1/2, Cel 1/2/3, Cel 6 or Cel E3/B5, or a stability region from one of the defined full-length sequences, or functional equivalents. Cel B5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid A1011 to P1424 or K1425 or N1426, and Cel B4/5 extends from amino acid X33 to D481, Cel E1/2 extends from Y39 to G812, Cel E1/2 EL/2/3 extends from Y39 to G812, Cel E1/2 extends from y39 to G812, Cel E3/B5 is the sequence shown in A4V13493; Cel E3/B5 is the sequence shown in A4V13493; Cel E3/B5 is shown in A4V13494. The new carying, backstaining or graying, for use on cellulosic materials including octton-containing fabrics. They are especially useful for preventing redeposition of colorant during stonewashing, and for preventing redeposition of colorant during stonewashing, and for truncated enzymes show reduced redeposition of dye compared to using non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim.
                                                                                                                               Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel E1/2; Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing; cotton-containing fabric; stonewashing; ss.
                                                                                                                                                                                                                                                                                                                                                                                                       Gibbs MD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 6415 BP; 2067 A; 1082 C; 1689 G; 1541 T; 0 U; 36 Other;
                                                                                                                                                                                                                                                                                                                                                                                                       Farrington GK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6415
38
14
32
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                               DNA sequence encoding truncated cellulases.
                                                                                                                                                                                                                                                                                                                                                                                                       Daniels RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Page 24-25; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                    (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 :runcated cellulase compositions
BP
AAX55662 standard; DNA; 6415
                                                                                                                                                                                                                                                                                                      98EP-00810919.
                                                                                                                                                                                                                                                                                                                                     97US-00932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.05e-13
                                                                                                                                                                                                                                                                                                                                                                                                     , Bergquist PL, Williams DP;
                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 194.50
61.18%
44.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-315403/27.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAY13493
                                                                                                                                                                                                                                                                                                      15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                       19-SEP-1997;
                                                                                                                                                                                                    Unidentified
                                                                30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                     EP921188-A2
                                                                                                                                                                                                                                                                     09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                         Anderson P,
                                                                                                                                                                                                                                                                                                                                                                                                                       Morgan H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ..
9
```

```
WPI; 1999-315403/27.
P-PSDB; AAY13492.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-SEP-1997;
                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-JUL-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EP921188-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                          84
                                                                                                                                                                                                                                                                                                                                                                                                              2902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX55661;
                                                                                                                                                                        Query Match:
                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                            Score:
88888888888
                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                          8
                                                   2724 TGGTTTAAGATAGTGAATGGAGGCAGCAGTGTTGATCTTAGCAGGGTTAAGATAAGA 2783
                                                                                                   2784 TACTGGTACACAGTGGAGGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2840
                                                                                                                                          2664 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACACACGTTCTATAAGGCCG 2723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic origin, where the cellulase active protein consists of the
                                                                                                                                                                                                                                                                                                         Active cellulase protein; alkalinophilic; textile processing; proteinase; detergent additive; stonewashed appearance; cotton-containing denim; CelB5; thermophilic; commercial detergent; celE gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and
                                        43
                                                                                63
                                                                                                                       83
23
                                                                                TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                      24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                       MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gibbs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P, Daniels R,
                                                                                                                                                                                                                                                                                                                                                                                                                                  encoding E1/2 protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  encoding El protein"
                                                                                                                                                                                                                                                                                                                                                                                                    "CelE protein'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bergquist
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Col 37-44; 61pp; English.
                                                                                                                                                                                                                                                                                      Active cellulase protein, celE gene
                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag= c /product= "DNA e 748. .2076 /*tag= b /product= "DNA e
                                                                                                                                                                                   GATTATTÁCCTGGAG 2915
                                                                                                                                                                                                                           AAD26526 standard; DNA; 6416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-00136574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97US-00932571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Anderson P,
                                                                                                                                                                AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                           /rcag= a
/product= '
                                                                                                                                                                                                                                                                                                                                                                                                              .2538
                                                                                                                                                                                                                                                                                                                                                                                534. .5889
                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Farrington GK, Andersor
Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       containing fabrics, is
elevated temperatures
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2002-081780/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-081780/
P-PSDB; AAE16324.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-AUG-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-SEP-1997;
                                                                                                                                                                                                                                                                                                                                                  Unidentified
                                                                                                                                                                                                                                                                   26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US6294366-B1
                                                                                                                                          2841
                                                                                44
                                                                                                                                                                                  2901
                                                                                                                        64
                                                                                                                                                                84
                                                                                                                                                                                                                                               AAD26526;
                                                                                                                                                                                                       RESULT 10
                                                                                                                                                                                                                 AAD26526
                g
                                                      g
                                                                                                  g
                                                                                                                                         셤
                                                                                                                                                                                  셤
                                                                                                                                                                                                                            à
                                                                             ò
                                                                                                                       ઠ્ઠ
                                                                                                                                                               ò
```

```
265 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACCAGGTTCTATAAGGCCG 2724
celBS amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as deterement additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing knits or for imparting stronwashed appearance to cotton-containing denims. The present proteins are stable under condition of alkaline pH and elevated temperatures, thus suitable for textile processing and in commercial detergents. The present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 GlyvalLysvalGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E1/2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gibbs MD;
                                                                                                                                                                                                                                                     Sequence 6416 BP; 2068 A; 1082 C; 1689 G; 1541 T; 0 U; 36 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cellulase; proteinase; truncated; Cel B5; Cel B4/5; Cel E1; Cel
Cel 1/2/3; Cel 6; Cel E3/B5; laundry detergent; stain-removing;
cotton-containing fabric; stonewashing; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ONA sequence encoding truncated cellulases Cel B4/5 and Cel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Farrington GK,
                                                                                                                                                                                                                                                                                                                            6416
38
14
32
                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Daniels RM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      x AAD26526 (1-6416)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CLRN ) CLARIANT FINANCE BVI LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAX55661 standard; DNA; 11706 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98EP-00810919.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   97US-00932571.
                                                                                                                                                                                                                                                                                                                         2.05e-13
194.50
61.18%
44.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Anderson P, Bergquist PL,
Morgan H, Williams DP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                          sequence is celE gene
                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-5 (1-88)
```

/product= "DNA encoding B4/5 protein"

9255. .10526

misc_feature

US6294366-B1 25-SEP-2001

/*tag=

protein'

/product= "CelB 8601. .10532

/*tag=

misc_feature

Location/Qualifiers

.10535

Key

/*tag=

/product= "DNA encoding B/5 protein"

```
The invention relates to a recombinant cellulase active protein free of proteinases of native thermophilic and alkalinophilic origin, comprising the truncated sequences Cel B5, Cel B4/5, Cel E1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 1/2, Cel 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ||||||||:::|||
TACTGGTACACAGTGGATGATGACAAGCCACAGAGTGCGGTA---TGTGACTGGGGCACAG 4214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCGAGCACAAGTTCTATAAGGCCG 4097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              43
          New truncated cellulase proteins, useful in detergents and for producing 'stonewashed' denim.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 11706 BP; 3828 A; 1994 C; 2994 G; 2890 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111706
37
16
31
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x AAX55661 (1-11706)
                                                                                                   Disclosure; Page 20-23; 65pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAD26525 standard; DNA; 11707 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     truncated cellulase compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GATTATTACTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.82e-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                192.50
62.35%
43.53%
41.04%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
```

New cellulase active protein, useful in textile processing or commercial detergents, e.g. for improving the feel or appearance of cotton-containing fabrics, is stable under conditions of alkaline pH and

Gibbs MD;

Bergquist P, Daniels R,

Anderson P,

Farrington GK, Anderso Morgan H, Williams DP;

WPI; 2002-081780/11.

P-PSDB; AAE16323

(CLRN) CLARIANT FINANCE BVI LTD

98US-00136574. 97US-00932571.

19-AUG-1998; 19-SEP-1997;

```
The present invention relates to a cellulase active protein, which is substantially free of proteinases of native thermophilic and alkalinophilic oxigin, where the cellulase active protein consists of the CelBs amino acid sequence. The cellulase active protein is useful for treating cellulosic materials including cotton-containing fabrics, as detergent additives. The cellulase active protein is also useful for improving the feel and/or appearance of cotton-containing fabrics, for removing surface fibers from cotton-containing whits or for imparting stonewashed appearance to cotton-containing whits or for imparting sure stable under condition of alkaline pH and elevated temperatures, thus sequence is celB gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4038 GGTTTGAAGGTACTATACAAGAACAATGAGACAAGTGCGAGCACAAGTTCTATAAGGCCG 4097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TACTGGTACACAGTGGTGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 4214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
:::|||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 11707 BP; 3827 A; 1995 C; 2994 G; 2891 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-5 (1-88) x AAD26525 (1-11707)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Col 27-38; 61pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.83e-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                192.50
62.35%
43.53%
41.04%
                                                                                                                                                                                                                                                                                                                                                                                                                                elevated temperatures.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

proteinase; denim;

Active cellulase protein; alkalinophilic; textile processing; detergent additive; stonewashed appearance; cotton-containing CelB5; thermophilic; commercial detergent; celB gene; ds.

Unidentified

Active cellulase protein, celB gene.

26-MAR-2002 (first entry)

AAD26525;

63

83

```
991 GGTATTTCTGTACAATACAGAGGAGGATGGGAGTATGAACAGCAACCAATCGGTCCG 1050
                                                                1051 CAGCITICAATAAAAAATAACGGCAATACCACGGITIGATITAAAAGATGTCACTGCCCGT 1110
                                                                                                                                                                  1111 TACTGGTATAACGCGAAAACAAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 1164
                                                                                             44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                      24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValArg
                                                                                                                                                                                                                                                                                                                                                                     CMCase gene with mutated signal sequence to enhance hydrophobicity.
                                                                                                                                                                                                                                                                                                                                                                                                Protein surface-displayed; genetic carrier; spore; monoclonal v
high-throughput screening; antibody; vaccine; enzyme; gene; ds.
                                                                                                                                                                                                                                                                                      AAL41025 standard; DNA; 1488 BP
                                                                                                                                                                                                                                 1225 GATACCTATCTGGAA 1239
                                                                                                                                                                                                       84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                             11-OCT-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                           Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pan JG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Key
                                                                                                                                                                                                                                                                                         CCCCCCCCCCCX8X4444X88X4X8X4X8X4X4444X8XXX4X4X8X
                                                                                                                                                                                                                               q
                                       à
                                                              셤
                                                                                             ò
                                                                                                                     g
                                                                                                                                                                         셤
                                                                                                                                                                                                       8
The invention relates to methods for preparing a protein surfacedisplayed on a genetic carrier, improving a protein or isolating a bubstance in a mixture comprising transforming a host cell harbouring a genetic carrier consisting of a spore or virus with the vector library. The method is useful for surface display of proteins, particularly for displaying a protein on the surface of e.g. spores, for improving a protein, or for isolating a substance. These methods are useful in obtaining monoclonal variants from a large library, or in high-throughput screening of antibodies for use in therapy (e.g. as vaccines), diagnosis of analysis. This polymucleotide sequence represents the CMCase gene with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Methods for surface display of proteins, e.g. displaying a protein on the surface of spores, improving a protein or isolating a substance, by transforming a host cell harboring a genetic carrier with spore or virus
                                                                                                                                                                                                                   Protein surface-displayed; genetic carrier; spore; monoclonal variant; high-throughput screening; antibody; vaccine; enzyme; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       or analysis. This polynucleotide sequence represents the CM additional His encoding sequences relating to the invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP; 493 A; 293 C; 341 G; 307 T; 0 U; 0 Other;
                                                                                                                                                                                          CMCase gene with additional His encoding sequences.
                                                                                                                                                                                                                                                                                                                                "CMCase protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 5; Page 109-113; 118pp; English.
                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                         AAL41028 standard; DNA; 1434 BP
                                          ||| ||||||:::
GATTATTACTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                                                                 15-JAN-2002; 2002WO-KR000059
                                                                                                                                                                                                                                                                                                                                                                                                                                          15-JAN-2001; 2001KR-00002156
                            84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Choi SK, Jung HC;
                                                                                                                                                                                                                                                                                                                  /*tag= a
/product=
                                                                                                                                                              11-OCT-2002 (first entry)
                                                                                                                                                                                                                                                                                                       .1434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GENO-) GENOFOCUS CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            with the vector library
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2002-590661/63.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-PSDB; AAO22444.
                                                                                                                                                                                                                                                                                                                                                           WO200255561-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1434
                                                                                                                                                                                                                                                            Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                      18-JUL-2002
                                                    4275
                                                                                                                                    AAL41028;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pan JG,
                                                                             RESULT 13
AAL41028
ID AAL41
                                                                                                                                                                                                                                                                                       Key
                        ò
                                                 g
```

variant;

```
The invention relates to methods for preparing a protein surfacedisplayed on a genetic carrier, improving a protein or isolating a substance in a mixture comprising transforming a host cell harbouring a genetic carrier consisting of a spore or virus with the vector library. The method is useful for surface display of proteins, particularly for displaying a protein on the surface of e.g. spores, for improving a protein, or for isolating a substance. These methods are useful in obtaining monoclonal variants from a large library, or in high-throughput screening of antibodies for use in therapy (e.g. as vaccines), diagnosis or analysis. This polymucleotide sequence represents the CMCase gane with a mutated signal sequence to enhance its hydrophobicity relating to the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Methods for surface display of proteins, e.g. displaying a protein on the surface of spores, improving a protein or isolating a substance, by transforming a host cell harboring a genetic carrier with spore or virus with the vector library.
Location/Qualifiers
1. .1491
/*tag= a
/product= "CMCase protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example 4; Page 98-103; 118pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JAN-2001; 2001KR-00002156
                                                                                                                                                                                                                                                                                                                                                                                              15-JAN-2002; 2002WO-KR000059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Choi SK, Jung HC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GENO-) GENOFOCUS CO LTD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-590661/63.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AA022443
                                                                                                                                                                                                                         WO200255561-A1
                                                                                                                                                                                                                                                                                                              18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                invention
```

4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23

Gaps:

US-09-917-376-5 (1-88) x AAL41028 (1-1434)

144 134 18 12 12 14 14

Length: Matches: Conservative: Mismatches: Indels:

1.94e-13

.. 80 ..

188.00 61.18% 40.00% 40.09%

Percent Similarity: Best Local Similarity:

Query Match: DB:

```
Choi S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                             Pan J,
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                     1108 CAGCTTCAAATAAAAATAACGGCAATACCACGGTTGATTTAAAAGATGTCACTGCCCGT 1167
                                                                                                                                                                                                                                                                                                                                                                                                   ||||||||||
TACTGGTATAAACGCGAAAAAAAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 1221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      :::|||||||||||||::
CTTGGATGCGGCAATGTGACATACAAGTTTGTGACGTTGCATAAACCAAAGCAAGGTGCA 1281
                                                                                                                                                                                                                                                                                                                                                                          TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                   83
                                                                                                                                                                                                                                                                                                       43
                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                     GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product= "Bacillus subtilis spore coat protein cotG-carboxymethyl cellulase (CMCase) fusion protein"
/note= "This sequence lacks a start codon"
461. .1045
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Spore coat protein; spore surface; delivery system; bioconversion; recombinant live vaccine; resistance property; protein microarray; whole cell absorber; secretion signal; target signal; cell membrane; surface display motif; core; CMCase; carboxymethyl cellulase; gene; caboxymethyl cellulase; gene; carboxymethyl cellulase; cellulase; cellulase; carboxymethyl cellulase; cellulase; cellulase; carboxymethyl cellulase; cellulase; cellulase; cellulase; carboxymethyl cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase; cellulase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacillus subtilis spore coat protein cotG-CMCase fusion gene.
              501 A; 297 C; 357 G; 333 T; 0 U; 0 Other;
                                                                   1488
148
118
2
                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1085. 2491
/*tag= e
/note= "CMCase structural gene"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               *tag= c
note= "CotG structural gene"
                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     promoter"
                                                                                                                                                                                            x AAL41025 (1-1488)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1046. .1084
/*tag= d
/note= "Linker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GATACCTATCTGGAA 1296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /*tag= a
/note= "CotG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABK53202 standard; DNA; 2510
                                                               2.04e-13
188.00
61.18%
40.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161. .2494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        *tag=
             BP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus subtilis.
                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                            US-09-917-376-5 (1-88)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W0200246388-A1
              Sequence 1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                1168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1282
                                                                                                                                                                                                                                                                                                     24
                                                                                                                                                                                                                                                                                                                                                                                                                                                   64
                                                                                                                                                                                                                                                                                                                                                                            44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABK53202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chimeric.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                promoter
                                                                                                                                      Query Match:
DB:
                                                                 No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          임
```

```
The present invention relates to a new method for displaying a protein of interest on a spore surface. The method of the invention involves containing a vector comparising a gene construct containing a gene encoding a spore coat protein and a gene encoding a protein of interest, a surface of a spore of the host cell, and recovering the combinant live vaccine. The method is also useful for producing a delivery system for recombinant live vaccine. The method is also useful for improving a sporte. The method is useful for producing a delivery spore, in combinant live vaccine. The method is also useful for improving a protein of interest using a resistance property of spore, in conversion reactions, for preparing protein microarray, for producing an antibody to antigen in vertebrates, and for preparing a whole cell carboxine. The spore coat proteins used circumvent a necessity for passage across cell membrane, so that they do not need secretion signal and carry a surface display of protein, such as beta-galactosidase, in corderly fashion, which is difficult to pass across cell membrane. The protein cotd-carboxymethyl cellulase (CMCase) fusion protein of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||||||||::
2171 TACTGGTATAACGCGAAAACAAGGCCAAAACGTT-----GACTGTGACTACGCGCAG 2224
                                                                                                                                                                                                                                                                                                                                        Displaying a protein of interest on spore surface, by transforming a host cell with a vector comprising a gene encoding a spore coat protein and a gene encoding a protein of interest, useful for preparing protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               521 G; 592 T; 0 U; 0 Other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2510
34
18
31
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x ABK53202 (1-2510)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example 4; Page 145; 158pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 2510 BP; 869 A; 528 C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2285 GATACCTATCTGGAA 2299
07-DEC-2001; 2001WO-KR002124.
                                                             08-DEC-2000; 2000KR-00074835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3.97e-13
188.00
61.18%
40.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84 AspThrTyrLeuGln 88
                                                                                                                       (GENO-) GENOFOCUS CO LTD
                                                                                                                                                                                                                                                 WPI; 2002-463830/49.
P-PSDB; AAU98063.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
```

Search completed: May 12, 2004, 05:39:39 Job time: 137.068 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

```
May 12, 2004, 04:33:52 ; Search time 826.547 Seconds (without alignments) 3179.336 Million cell updates/sec
                                                                                                                                          469
1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ 88
OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                          Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                         US-09-917-376-5
                                                                                                                                                                                                         BLOSUM62
                                                                                                                                            Perfect score:
                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                   Sequence:
                                        Run on:
```

Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 2000000000

55026578

27513289 seqs, 14931090276 residues

Searched:

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-Q=(Q=10_21/USPFD spool/1809917376/runat 11052004 114535 28249/app query.fasta_1.1429
-Q=(Q=10_21/USPFD ospool/1809917376/runat 11052004 114535 28249/app query.fasta_1.1429
-DB=EST -QFWT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bLoomm62 -TRANS=human140.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXIEN=200000000
-USER-USO9917376 @GGN 1 1.5334 @runat 11052004 114535 28249 -NCFU=6 -ICPU=3
-NO PMARP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gssl:* 1: em estba:*
2: em_esthum:*
3: em_estin:*
4: em_esten:*
5: em_esten:*
6: em_estpl:*
7: em_estpl:*
8: em_htc:*
9: gb_estl:*
10: gb_estl:*
10: gb_estl:* em_gss_hum:* em_gss_inv:* em_gss_vrt:* em_gss_fun:* em_gss_mam:* em_gss_mus:* em_gss_pln:* em_estfun:* em_estom:* gb_est3:* gb_est4:* gb_est5:* EST: * Database :

gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

S
r-3
-
м
~
-
ď
_
2:
₽
~
ъ
E/O

ription	15	7000	10000	2000		52891	26529	2	37649	39392	16323	18457	0	741	₩.	1128	1562	1464	5709	1843	5729	BI664037 603289338	2908	3980	3084	1904	5850	3232	38 C64738 Yu	2108 laf13d0	56	CH230	353 PUFR	785	6375 dads	5 AL799805	3120 EST375	543 AL898643	716 10001853	1312 EST-77-	376 AL96257	3602 tigr-	5493 AL846493		35313 AT31857.	
_		000430	20,4000	000000	2002/00 31750044	BE36289	CD22652	BE3628	CF48764	BG73939	BZ61632	BI2184	BZ5	CG2707	CG2707	BF5131	CK2315	AW4914	CA5957	BZ8948	CF0767	BI664037	BF8759	CB2939	CB2930	BG33490		CF07823	C64738	CF6221	BX2094		CG0778	CG0778	BG8163	AL79980	AW9631	AL89864	BZ4247	CF754312	AL962576	CE2536	AL846493		BG63531	
g E		* *	# 0	9 4	376	, ,	41 1	57 1	76 1	91 1	24 2	14 1	05 2	34 2	01 2	657 1	58 1	65 1	13 1	50 2	02 1	653 12	61 1	16 1	819 1	039 1	37	57 1	09	93 1	76 2	61 2	94 2	24 2	80 1	20	33	44 9	563 28	29	79	01	40	48	19	`
거스	1	•			1.01										16.2								15.7				15.7					15.6			٠											•
		٥	, r	ċr		76	16	16	16	16	94	76	16	16	7	'n	ď.	4.	74.5	4.	74	7	ë	ë.	73.5	e.	'n.	73	73	73	73	73	73	7	'n	'n	ς.	ď.	72.5	ď	ά.	2	2	2		;
Result No.		٦ ،	7 C	า <	4 U	י נ	, ,	- 00	0	10	c 11				15		-	٦	c 19	7	21	22	c 23	24	c 25	56	c 27	28	29	30	31	32	c 33	34	c 32	m	37	'n	c 39		c 41	42	4	C 44	4	}

ALIGNMENTS

EST 08-SEP-2003 EST.
Sorghum bicolor (sorghum)
Sorghum bicolor
Sorghum bicolor
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Bukaryota; Viridiplantae; Liliopsida; Poales; Poaceae; PACCAD
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum. S84 bp mRNA linear EE POLL 65_H07.bl A002 Pollen Sorghum bicolor cDNA clone POLL 65_H07_A002 3', mRNA sequence.
CF480372 CF480372.1 GI:34509241 VERSION KEYWORDS SOURCE ORGANISM LOCUS DEFINITION ACCESSION RESULT 1 CF480372

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; saterida; lamida; Solanales; Solanaceae; Solanum.

I (bases 1 to 754)

Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukoianov, A., Rangel, P., Haberlach, G.T., Karamycheva, S.A., Tsai, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B. Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, incompatible Interaction (2002)

Unpublished (2002)

Contact: Robin Buell

The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics in the contact of through URL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /nocessign for the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the co
                                                                                                                                                                                                                                                                BQ047524 10-MAR-2003
EST596642 P. infestans-challenged potato leaf, incompatible
reaction Solanum tuberosum cDNA clone BPL117P18 5' end, mRNA
   213 GCTTGTGTGAAGTGTGGTGACGTTGGTTTTTGGAGACGGTGTCTTGGTCTGTTGGTCCG 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7 ValGlnTyrLysAenAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGln 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db xref="taxon:4113"
/clone="BPL117P18"
/tissue type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/clone_lib="P. infestans-challenged potato leaf,
incompatible reaction"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         754
29
114
33
4
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Solanum tuberosum"
/mol type="mRNA"
/cultivar="Kennebec"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           http://genome.arizona.edu/orders/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x BQ047524 (1-754)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Seq primer: T3.
    Location/Qualifiers
    1. .754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Solanum tuberosum (potato)
                                                                                                        273 AGCTTGCAGTCTGGTGAT 290
                                                                   84
                                                                                                                                                                                                                                                                                                                                                                                                 BQ047524
BQ047524.1 GI:19821510
                                                                   79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27.9
79.50
38.05$
25.66$
16.95$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Solanum tuberosum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local S.
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
VERSION
                                                                                                                                                                                                    RESULT 2
BQ047524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                   g
                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /db xref="taxon:4558"
/clone="PoL1 65_H07_A002"
/lab host="MloB-T1 phage-resistant E. coli"
/lab host="MloB-T1 phage-resistant E. coli"
/clone lib="PoL1 65_H07_A002"
/clone lib="PoL1 65_H07_A002"
/clone lib="Pol1en; Vector: pME18S-FL3; Site 1: XhoI;
/clone lib="pol1en; Vector: pME18S-FL3; Site 1: XhoI;
/clone pollen: Pollen vacuolated-vacuolated stage of development. Pollen was harvested from greenhouse-grown panicles of sorghum line BTx623. Panicles were removed from the flag leaf prior to emergence, when no detectable amylase is present in pollen of male-fertile lines. This stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACGATGTG). XhoI excises the CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Email: mmpratt@ugg.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science, plant material and RNA prepared at Texas A & W University,
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq prime: Sug3-14 (TAGTCTAGCGGCGGACC)
                                     Liang, C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105 -------CGCGCCGGACAACCTGAAACTTCCAAACGATGGTGCAAGAAA 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCCTGTGCTGGCCCGGGC 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78
                           Cordonnier-Pratt, M. M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J., Lucas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                    Est database from Sorghum: pollen
Unpublished (2003)
Other ESTS: POLL 65_H07.gl A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        584
24
15
37
10
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |. .584
|organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-5 (1-88) x CF480372 (1-584)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /mol_type="mRNA"
/cultivar="BTx623"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12.7
81.00
45.35%
27.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                           TITLE
JOURNAL
COMMENT
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

셤

ð 셤 ઠે

ò

ò

셤

ઠે

```
/clone="UVCRCS01" (55012" / 1.1 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.2 / 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CB293085 UCRCS01_05cb12_g1 Washington Navel orange cold acclimated flavedo & albedo cDNA library Citrus sinensis cDNA clone UCRCS01_05cb12, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 (bases 1 to 768)
Close,T.J., Collura,K., Fenton,R.D., Feuerbacher,O., Kim,H.R.,
Kudrna,D., Wanamaker,S., Wing,R. and Yu,Y. and Development of EST Resources and New Genetic Markers for California
                                                                                                                                                                      268
                                                                                                                                                                                                                                                                                                                                       269 AAGGGGGAGCCTCCGAAGTTTCTTGGTTATGAACATTTCTCCACTCGCAAATGGGGGTTCA 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots; rosids, eurosids II, Sapindales, Rutaceae, Citrus.
GTGCCTTATCAAGATCCAGGAAGT----CCGGGTAGTAATCTAATATCTCCTGGTTCAGTA 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             329 AGGGTTGGCTCAGGATCATTGACACCAAGTGGCTGGGGCTCTAGGCTTAGGTTCTGGAACT 388
                                                                                                                                                                                                                                                         47
                                                                                                                                                                                                                                                                                                                                                                                                                               54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
                                                                                                                              27 LeuValAsnThrGlySerSerSer------
                                                                                                                                                                                                                                                         ---ValAspLeuSerThrValThrValArgTyrTrpPheThr
                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GlySerSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55 LeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           75 SerValAsnProAlaThrProThrAlaAspThrTyrLeu 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /mol_type="mRNA"
/cultivar="Washington navel"
/db_xref="taxon:2711"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'organism="Citrus sinensis"
                                                                                                                                                                                                                                                                                                                                                                                                                          48 ArgAspGly-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Fax: 9097874437
Email: timothy.close@ucr.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CB293085.1 GI:28618542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unpublished (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Citrus sinensis
Citrus sinensis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tel: 9097873318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seq primer: T3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CB293085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Citrus
                                                                                                                                                                      209
                                                                                                                                                                                                                                                    35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CB293085
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
```

```
Sci. (1996) 64:809-814. Poly(A) RNA was purified, a cDNA library was made, and 1 million primary lambda CDNA clones were in vivo excised to give a population of pBluescript SK(-) phagemids. All steps to this point were performed in the TJ Close lab at the University of California, Riverside (Fenton). Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3530 at the Arizona Genomics Institute, University of Arizona (Collura, Feuerbacher, Kim, Kudrna, Wing, Yu). Chromatogram files were transmitted to UC Riverside (by Yu), then processed at UC Riverside (by Wanamaker) using the HarvEST pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BQ722991

AGENCOURT 8103733 Lupski sympathetic_trunk Homo sapiens cDNA clone IMAGE:6191052 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  612 GGCGGCATCAAC-----AAGAACAACAACAGCAGCAGCAATAATAATAAAAAGTAAT 662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              663 AATAATAGCGACGTTGTTACTAGTGGTGGCGTTAGGTGCAGTGCGCGTGACG -- 719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. I (bases 1 to 906)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 GlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLys 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Email: cgapberfemail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llhl.gov.row: n column: 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23 ProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           768
20
10
15
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 720 -----TITITGGAAGGCGAGGAGCAGCAGT 746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43 ArgTyrTrpPheThrArgAspGlyGlySerSerThr 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /mol_type="mRNA"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5 (1-88) x CB293085 (1-768)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 High quality sequence stop: E
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BQ722991
BQ722991.1 GI:21861888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37.9
78.50
57.69%
38.46%
16.74%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 4
BQ722991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

```
Insert Length: 1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28.5
76.00
43.42$
26.32$
16.20$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pratt, L.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
BE362891
                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                              1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life
                                         EST 11-MAY-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ::::::||| ||| ||| |||   |||::: |||   |||   ||:::::   |||   ||   ||:::::   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||   ||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi, Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
I (bases I to 376)
Hillier,i., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,N., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Travaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7 ValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGln 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27 LeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrpPhe 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 zk61f10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone IMAGE:487339 3' similar to gb:L10240 BASIGIN PRECURSOR (HUMAN);,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TCCCGG---GGCGCGCTGCCCCAGCTCCTACACCTGCAGCCAGTGGTGGA 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376 bp
clone="IMAGE:6191052"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-917-376-5 (1-88) x BQ722991 (1-906)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA043716.1 GI:1521725
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Technologies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73.1
77.00
49.15$
27.12$
16.42$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              and Marra, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA043716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA043716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97044478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8889549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
AA043716
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred.
```

ò 셤 ઠે 셤 8 원

```
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BE362891 606 bp mRNA linear EST 17-DEC-2001
DG1_90_C11.91_A002 Dark Grown 1 (DG1) Sorghum bicolor CDNA, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    178 CCCCTCCAGAGCTGCCCCCGGATGCACAGACGAGGGGGGCTTGGGAGTGACTTGAGGC 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sorghum bicolor
Sukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------LeuValTyrAsn 58
                                                                                                                                                                                                                                                         /sex="female"
/dev_stage="adult"
/lab_host="bHiOB"
/clone lib="Soares pregnant_uterus NbHPU"
/note="Grgan: uterus; Vector: pT7T3-Pac; Site 1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17 GlyAspAsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        37 LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThr----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 606)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M.
                                                                                                                                                                                                                                                                                                                                                                                                                        TGTGACGGGGTCGCCCTCGGGTGTGGCAAGTGAAGTCCTCTGTGGGC 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59 CysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
On Jul 20, 2000 this sequence version replaced gi:5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            376
20
13
23
20
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Std Error: 0.00
             Seg primer: -40M13 fwd. from Amersham
High quality sequence stop: 212.
Location/Qualifiers
                                                                                                                     /organism="Homo sapiens"
mol_type="mRNN"
/mol_type="mRNN"
/db_xref="GDB:3761053"
/db_xref="taxon:9606"
/clone="IMAGE:487339"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-5 (1-88) x AA043716 (1-376)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BE362891
BE362891.2 GI:15724434
```

```
BE362820 657 bp mRNA linear EST 20-JUL-2000 DG1 89_C02.g2_A002 Dark Grown 1 (DG1) Sorghum bicolor CDNA, mRNA sequence.
                                                                                                                                                          Email: mmprattedga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /clome_lib="Callus culture/cell suspension"
/note="Vector: pMRI8S-F13; Site_1: Xho1; Site_2: Xho1; Th
library was prepared from a mixture of polyA+ RNA from
callus culture tissue and cells in suspension culture.
Double-stranded cDNA was cloned unidirectionally into
different DraII sites of the pMRI8S-F13 vector (5-prime
DraII site is CACTGTGTG, 3-prime DraIII site is
CACCATGTGT, Xho1 excises the cDNA insert."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       149 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCCGGAGGAGGAGCCTGTGCTGGCCCGGGG 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                           Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 543 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:4558" .
/clone="CCC1 46 E07 A007"
/lab_host="DH10B-T1 phage-resistant E. coli"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               641
23
15
38
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                              Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
POLYA=Yes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                          organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
         Other ESTs: CCC1 46 E07.gl A007
Contact: Cordonnier-Pratt MM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-5 (1-88) x CD226529 (1-641)
                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /mol_type="mRNA"
/cultivar="RTx430"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            377 AGCTTGCAGTCTGGTGAT 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GI:9304377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            59.9
76.00
44.19%
26.74%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .641
                                                                                                                                                                                                                                                                                                                                                  exclude polyA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BE362820
BE362820.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
BE362820
                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11 bp mRNA linear EST 21-MAY-2003 culture/cell suspension Sorghum bicolor
                                                                                              Sequences have been trimmed to exclude PolyA, vector and regions below Phrea quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTMix or 77 sequencing primer, are presented as the reverse complement.
                                                                                                                                                                                                                                                                                                                                                                                     /clone lib="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Inabda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------CGCGCCGGGACAGAACCCTGAAACTTCCAAACGATCGGTGCAAGGAAA 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GCTTGTGTGAAGTGTGACGTTGGTTCTTTGGAGACGGTGTCTTGCTCTTGTTTGGTCGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cordonnier-Pratt.M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Rathore,K.,
Esstman,A. and Pratt,L.H.
An EST database from Sorghum: callus culture and cell suspension
Unpublished (2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCC1_46_E07.bl_A007_Callus culture/cell suspens
cDNA_clone CCCI_46_E07_A007_3', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           606
23
115
110
110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                           organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                     Seq primer: PolyTMix
High quality sequence start: 15
High quality sequence stop: 594
POLYANO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  641 bp
                                                                                                                                                                                                                                                                                                                                             /mol_type="mRNA"
/db_xref="taxon:4558"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5 (1-88) x BE362891 (1-606)
                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sorghum bicolor (sorghum)
Sorghum bicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGCTTGCAGTCTGGTGAT 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AlaThrProThrAlaAsp 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CD226529.1 GI:30969963
                                                                           Email: mmpratt@uga.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55.4
76.00
44.19%
26.74%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (bases 1 to 641)
                                                                                                                                                                                                                                                                                                  909.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CD226529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CD226529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
DB:
                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
CD226529
                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
```

ઠે 요 8 g 43

61

g

ઠ

요

δ

g

ठ

```
/note="Organ: Pollen; Vector: pME18S-F13; Site 1: Xho1; Site 2: Xho1; The library was prepared from poly4+ RNA from pollen at the late vacuolated-vacuolated stage of development. Pollen was harvested from greenhouse-grown panicles of sorghum line BTx623. Panicles were removed from the flag leaf prior to emergence, when no detectable amylase is present in pollen of male-fertile lines. This stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different Drall sites of the pME18S-F13 vector (5-prime Drall site is CACTGTGTG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genter, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & W University,
Science; plant material and RNA prepared at Texas A & W University,
Sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
                                                                                                                           Bukaryotta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 6/76)

Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C. Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J., Ducas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and Pratt, L.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -prime DraIII site is CACCATGTG). XhoI excises the cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 543 0210
Pax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /mol_type="mRNA"
/cultivar="BTx623"
/db_xref="texon:4558"
/clone="POL1 45 C08 A002"
/lab host="DH108-T1 phage-resistant B./clone_lib="Pol1en"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seq primer: Sug3-14 (TAGTCTAGCGGCCGCGACC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                               EST database from Sorghum: pollen
                                                                                                                                                                                                                                                                                                                                                                                                                                    Unpublished (2003)
Other ESTs: POL1 45 C08.gl A002
Contact: Cordonnier-Pratt MM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5 (1-88) x CF487649 (1-676)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                           Sorghum bicolor (sorghum)
          GI:34516518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64.6
76.00
44.19%
26.74%
16.20%
                                                                                                         Sorghum bicolor
          CF487649.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLYA=Yes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                     KEYWORDS
SOURCE
                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
          VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EST 08-SEP-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /mol type="mana" | /mol type="mana" | /mol type="mana" | /mol type="mana" | /mol type="mana" | /mote="mana" | /mote="mana" | /mote="mana" | /mote="mana" | /mote="mana" | /mote | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mota | /mot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoldeae, Andropogoneae, Sorghum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       204 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCCTGTGCTGGCCCGGGC 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg--- 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44 ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly
                                                                                                                                                                                                                                                                                                                                                                                                  Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
                                                                                                                                                                                                  1 (bases 1 to 657)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLI_45_C08.bl_A002 Pollen Sorghum bicolor cDNA clone POLI_45_C08_A002_3', mRNA sequence.
                                                                                                                                                                                                                                                                   Pratt, L. H.
An EST database from Sorghum: dark-grown seedlings Unpublished (2000)
Contact: Cordonnier-Pratt MM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          657
23
15
38
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="Sorghum bicolor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                dq 9/9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seg primer: PolyTMix
High quality sequence start: 89
High quality sequence stop: 651
PoLYA=No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (1-657)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                     Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      432 AGCTTGCAGTCTGGTGAT 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Email: mmpratt@uga.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x BE362820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44.19%
26.74%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62
76.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tel: 706 542 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .657
                                                                Sorghum bicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CF487649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                    ORGANISM
                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
CF487649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
```

g

ò

ò g 8 g ò

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (Bases 1 to 724)

2 Rabinowicz, P. D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)

Unpublished (2002)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8884

Fax: 516 367 8884

Fax: 516 367 8884

Calses: shotgun

Class: shotgun

Class: shotgun

Class: shotgun
                                                                                                                                                                                                                                                                                                                                                                                                                         124 bp DNA linear GSS 16-JAN-2003 ig58e08.gl WGS-ZmaysF (DHSa methyl filtered) Zea mays genomic clone 1958e08 5', genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /lab host="DBGs"
/clone lib="wGS-ZmaysF (DH5a methyl filtered)"
/clone=lib="wGS-ZmaysF (DH5a methyl filtered)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with Xbal and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size.
fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mpl9: .b/g reads in pUC19). The same ligation was transformed into DH5a."
                                                                                                                                                                                                        ------CGCGCCGGGACAGAACCTGGAACTTCCAAACGATCGGTGCAAGGAAA 372
                                                       44 ---TyrTrp----PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                     62 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            724
34
7
7
29
43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           High quality sequence stop: 728.
Location/Qualifiers
1. 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /db_xref="taxon:4577"
/clone="ig58e08"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-917-376-5 (1-88) x BZ616323 (1-724)
                                                                                                                                                                                                                                                                                                                         493 AGCTTGCAGTCTGGTGAT 510
                                                                                                                                                                                                                                                                                      84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BZ616323.1 GI:27766878
                                                                                                                                                                                                                                                                                   79 AlaThrProThrAlaAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            71
76.00
36.28
30.09
16.20
28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
  325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
BZ616323/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                  g
                                                                                                                                                                  ò
                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Sorghum bicolor"
/mol type="mRNA"
/db_xref="taxon:4558"
/clone lib="Embryo 1 (EM1)"
/note="Organ: Embryo germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector: Embryos germinated for 25 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector: Embryos germinated for 24 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 24 hr; Vector:
/note="Organ: Embryos germinated for 24 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germinated for 25 hr; Vector: Embryos germin
                                                                                                                                                                                                                                                                                                                                                                                                                         BG739392 691 bp mRNA linear EST 15-MAY-2001
EM1_82_D05.91_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
                                                                                                                                                                           Eukāryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 265 ATAGCAATCAGCCGAAACTCAAATGAATCCGGTCCGGGAGGAGCCTGTGCTGGCCGGGC 324
                                                                                 78
                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 ValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25 LeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg--- 43
                          ---TyrTrp---PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                        62 AlaAlaMetGlyCysGlyAsnIle-----ArgAlaSerPheGlySerValAsnPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pratt, L.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and
An EST database from Sorghum: developing embryos
Unpublished (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          691
23
15
38
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 691)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seq primer: PolyTMix
High quality sequence start: 50
High quality sequence stop: 691
PoLYA=No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-5 (1-88) x BG739392 (1-691)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: Cordonnier-Pratt MM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sorghum bicolor (sorghum)
                                                                                                                                                                                                                                                                                               442 AGCTTGCAGTCTGGTGAT 459
                                                                                                                                                                                                                                                      79 AlaThrProThrAlaAsp 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BG739392
BG739392.1 GI:14089081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66.6
76.00
44.19%
26.74%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tel: 706 542 1860
Fax: 706 583 0210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sorghum bicolor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similari
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment ?
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                            BG739392
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                           g
                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                     ò
                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

```
US-09-917-376-5 (1-88) x BI218457 (1-814)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-376-5 (1-88) x BZ545807 (1-905)
                                                                                                                                                   ----- 06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zea mays
Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
BZ545807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                ò
                                                                                                                                              원
                                                                                                                                                                                  ò
                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                           BI218457 814 bp mRNA linear EST 11-JUL-2001
602937951F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5101069 5',
                                   528
                                                                                               527 édédececeadergadeagergédadegadadecaceaciteargéritareadeceregeege 468
                                                                                                                                                               467 ACGAGGAAGTTCGTCAACTTCCCGGAGATCGTCGGAGCGGTCCAGAAGGCTGGGTTCGAG 408
                                                                                                                                                                                                                               407 GTCATCCCGATCGAGCCGCGTAGAGACCTCAGCGTGGAGGACTTCGCCCGGACCGTCGAC 348
                                                                                                                                                                                                                                                                                             Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
                                                                                                                                   36
                                                                                                                                                                                                  23
                                                                                                                                                                                                                                                               73
      -----AlaPro 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db.mcme="IMAGE:5101090"
/clone="IMAGE:5101069"
/lab_host="DH108 | The phage-resistant)"
/clone=lib="NCI CGAP_Li9"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1:9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP_Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: Ggapbs-remail.nih.gov
Tissue Procurement: Jeffrey B. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://mage.llnl.gov.f column: 14
                     17 Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn-----
                                                                                                                                                                                          37 ------LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSe
                                                                                                                                                                                                                                                             53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 (bases 1 to blt)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       814
28
11
29
19
                                                                                                                                                                                                                                                                                                                                                73 eGlySerValAsnProAlaThrProThrAlaAspThr 85
SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                               30 Thr-GlySerSerValAsp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       High quality sequence stop: 792
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mus musculus (house mouse)
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BI218457.1 GI:14671901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83.6
76.00
44.83%
32.18%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. .814
                                                                                                                                                                                                                                                                                                                                                                                                                                                             mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BI218457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                             316
 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                               BI218457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                              g
                                                                                            g
                                                                                                                                                             셤
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                             ద
   ð
                                                                 ò
                                                                                                                               ò
                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                           ò
```

```
CALR72TW ZM2 0.7 1.5 KB Zea mays genomic clone ZMMBMa0089L24, genomic survey sequence.
                                                                                                                                                                                                                          176
                                                                                                                                -----AGTTCCCGAGTGGACCTGAGGACCCTG--- 116
                                                                                                                                                                                                                                                                                                Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta; Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD clade, Panicoideae, Andropogoneae, Zea.
                                                                                                                                                                                29
                           22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                       60 Asp-TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAl 79
  21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 905)
Mitchaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Frager, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maze Genomics
Unpublished (2002)
Other_GSSS: OGALR72TC
                                                                                                                                                                                                          2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                              42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeu-----ValTyrAsnCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /clone lib="ZM2_0.7_1.5 KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic_DNA_library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9712 Medical Center Drive, Rockville, MD 20850, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        905
7 4 4 2 4 3 9 6 5 6 5 6 9 6 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 6 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .905
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="taxon:4577"
/clone="ZMMBMa0089L24"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                       225 TACGACGCTGCGATTACA 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                               79 aThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BZ545807
BZ545807.1 GI:27099699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seq primer: TR
Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           76.00
36.28%
30.09%
16.20%
```

---- 534

53

셤 ò g à 임 ò g

ò 셤 ò g

```
CGNIB3TV ZM 0.7_1.5_KB Zea mays genomic clone ZMMEMa0583E18, genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zea mays
Zea mays
Zea mays
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae, Andropogoneae; Zea.

(hases 1 to 1001)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Praser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Consortium for Maize Genomics
                                                                                                      624 GTCATCCCGATCGAGCCGCGTAGAGACCTCAGCGTGGAGGACTTCGCCCGGACCGTCGAC 565
                                                                                                                                                  30 Thr-GlySerSerSerValAsp----- 36
                                                                                                                                                                                                                                                                                  53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh 73
                    SerGlyGlyValLysValGlnTyrLysAsnAsnAspSer------AlaPro 16
                                                                                                                                                                                                                 37 ------LeuSerThrvalThrvalArgTyrTrpPheThrArgAspGlyGlySerSe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref="taxon:4577"
/clone="zwHMM0583B18"
/clone=lib="zm_0.7 1.5_KB"
/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 ]
methylation filtered genomic DNA library"
                                                                                17 Gly-------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1001
34
7
29
43
                                                                                                                                                                                                                                                                                                                   564 TCGTGCG-----ACGTGCTCATGGGCGCGCACGGGG-----
                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="genomic DNA"
/strain="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="Zea mays"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Class: sheared ends.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CG270748
CG270748.1 GI:34182889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished (2002)
Other_GSSs: OGWIB33TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1112
76.00
36.28
30.09
16.20
29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1, .1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seq primer: TF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
CG270748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
                                                                                ò
                                                                                                              g
                                                                                                                                                  ò
                                                                                                                                                                               셤
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                     ሯ
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                            OGWIB33TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBMa0583E18, genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 934)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Unpublished (2002)
147 GGCGCCCCAGCTGAGCCGCAGCCACCACCACGATGCTTATCAACCGTGGCCGC 206
                                                                                                                                                    267 GTCATCCCGATCGAGCCGCGTAGAGCCTCAGCGTGGAGGACTTCGCCCGGGACCGTCGAC 326
                                                                                                                                                                                                                                                                                                   327 TCGTGCG-----ACGTGCTCATGGGCGCACGGGG------357
                                                                    29
                                                                                                                                   36
                                                                                                                                                                                                37 ------LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSe 53
                                                                                                                                                                                                                                                                   73
                                                                17 Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn-----
                                                                                                                                                                                                                                                                   53 rThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPh
                                                                                                                                 30 Thr-GlyserSerValAsp------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref="taxon:4577"
/clone="zwMBWab581818"
/clone=b=zwMBWab581818"
/clone=b=zwm.0.7.1.5.KB"
/note=vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9712 Medical Center Drive, Rockville, MD 20850, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          934
7
7
29
43
                                                                                                                                                                                                                                                                                                                                                      73 eGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="genomic DNA"
/strain="B73"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         organism="Zea mays"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-5 (1-88) x CG270741 (1-934)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CG270741 GI:34182882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Other_GSS8: OGWIB33TV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seq primer: TR
Class: sheared ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          76.00
36.28%
30.09%
16.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. .934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zea mays
Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                    CG270741/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NIGINC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
```

```
.---- 688
                                              ||| |||||| ::: ||||||| 418 TCCGCAGCGTCACTCCCATACAAGAGGCGAAACACAACGCTGCGCACACCCC 477
                                                                                                      538 ACGAGGAAGTTCGTCAACTTCCCGGAGATCGTCGGAGGGCTCCAGAAGGCTGGGTTCGAG 597
                                                                                                                                                                                                                         598 GTCATCCCGATCGAGCCCGTAGAGACCTCAGCGTGGAGGACTTCGCCCGGACCGTCGAC 657
                         2 serglyglyvalLysvalGlnTyrLysAsnAsnAspSer-------AlaPro 16
                                                                                                                                           30 Thr-GlySerSerSerValAsp----- 36
                                                                                                                                                                                                    37 ------LeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSe 53
                                                                                                                                                                                                                                                              Gly------AspAsnGlnIleLysProGlyLeuGlnLeuValAsn----
                                                                                                                                                                                                                                                                                                                      658 rcgrgcg-----acgrgcrcargggcgcgcacggg----
US-09-917-376-5 (1-88) x CG270748 (1-1001)
                                                                                                                                                                                                                                                                                                                                                                                                Search completed: May 12, 2004, 11:39:02
Job time : 828.547 secs
                                                                                    11
                                                                                                                                                                      a
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                      g
                                                    d
                                                                                    ò
                                                                                                               g
                                                                                                                                              ò
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                         ò
```

Appli Appli

Appli Appli Appli Appli Appli Appli Appli Appli Appli

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Total number

Searched:

```
Fabrics Using Truncated
                                                                     Sequence
Sequence
Sequence
                                                                                                                    Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                        Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
COUNTY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: AURKOWIN
PRIOR APPLICATION DATA:
                       Sequence
Sequence
Sequence
                                                                                                                                                                    Sequence
                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                             Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                           Sequence
                                                                                                                                                                                             Sequence
                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                             Sequence
                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Containing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Treating Cellulose Containing
Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Morgan, must be positions and Methods for TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing
                                                                                                                                          US-09-006-636-5

US-09-006-632-5

US-09-277-716-21

US-09-375-714-5

US-09-609-161B-21

US-09-335-586-3

US-09-526-193A-21

US-09-453-702B-39

US-09-453-702B-39

US-09-453-702B-39

US-09-233-336A-20

US-09-233-356-20

US-09-233-356-20

US-09-233-752A-20

US-09-233-752A-20

US-09-233-752A-20

US-09-404-226-20

US-09-818-219B-19
                                                                                                                                                                                                                                                                                                                               -09-233-336A-19
                                                                                                                    US-08-330-394A-1
US-08-330-394A-3
                                                                                                                                                                                                                                                                                                                                                      -09-402-036-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
US-09-136-574A-46
; Sequence 46, Application US/09136574A
; Patent No. 6.94366
; GENERAL INPORMATION:
APPLICANT: Farrington, Graham
Anderson, Paige
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Bak, Mary E
                                                                                                                                                                    16.4
115.5
114.9
14.4
14.4
14.4
                                                                                                                                                                                                                    75.5
                                                                                                                                                                                                                                                       Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                 ; Search time 21.976 Seconds
  (without alignments)
2222.227 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 46, 1
Sequence 2, Al
Sequence 1, Al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 11, Sequence 9, A Sequence 11, Sequence 6, A Sequence 6, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 11,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 3,
Sequence 9,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                         1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                           nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-136-574A-46
US-09-136-574A-1
US-09-136-574A-1
US-07-862-588B-1
US-09-339-159B-3
US-09-138-956-9
US-09-198-955A-11
US-09-694-531-11
US-09-670-141-9
US-07-862-588B-6
US-07-862-588B-6
                                                                                                                                                                                                                                                                              of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                        682709 segs, 277475446 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                        Match 100%
first 45 summaries
                                                                                 May 12, 2004, 04:43:36
                                                                                                                                                                                           Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100
                                                                                                                                 US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6416
111707
2977
1483
1482
1482
1482
1624
1775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                    Listing
                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                            Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
```

1995.5 1994.5 192.5 1655 1655 1655 1655 1655 1655 120.5

Score

Result

Database

```
Containing Fabrics Using Truncated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2665 GGTGTGAAGGTACTGTACAAGAACAATGAGACAAGTGCGAGCACAGGTTCTATAAGGCCG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2785 TACTGGTACACAGTGGATGGTGACAAGCCACAGAGTGCGGTA---TGTGACTGGGCACAG 2841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
                                                                                                                                                                                                                                                                                                                                                                                                                              6416
38
14
32
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Treating Cellulose Containing Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.
TITLE OF INVENTION: Compositions and Methods for
                                                                     APPLICATION NUMBER: US 08/932,571
FILING DATE: September 19, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REPERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x US-09-136-574A-2 (1-6416)
     APPLICATION NUMBER: US/09/136,574A
                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                            NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Farrington, Graham K.
Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                     FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 1, Application US/09136574A; Patent No. 6294366; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID
                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 6416 base pairs
                                                                                                                                                                                                                                                                                                         TYPE: nucleic acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2902 GATTATTACCTGGAG 2916
                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88
                                                                                                                                                                                                                                                                                                                                                                                                                                    7.26e-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                194.50
61.18%
44.71%
41.47%
                                                                                                                                                                                                                                  TELEX: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84 AspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                             US-09-136-574A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . <u>අ</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 ATAAGGCCGTGGTTTAAGATAGTGAATGGAGCAGCAGCAGTGTTGATCTTAGCAGGGTT 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspAsrAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET: Spring House Corporate Center, P.O. Box 457 CITY: Spring House
                                                                                                                                                                                                                                                                                   2029
38
15
15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Williams, Diane P. TITLE OF INVENTION: Compositions and Methods for
                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: 1997US001/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x US-09-136-574A-46 (1-2029)
                                                                                                                                                                                               TOPOLOGY: linear SEQUENCE DESCRIPTION: SEQ ID NO: 46:
                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238 AGTGGAGCGGATTATTACCTGGAG 261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Farrington, Graham K.
                                                                                INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 2029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application US/09136574A Patent No. 6294366 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Daniels, Roy
Morgan, Hugh W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                       1.06e-16
                                                                                                                                                                                                                                                                                                       195.50
60.23%
43.18%
                                                                                                                                                                                                                                                                                                                                                          41.68%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE: PA
                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                  US-09-136-574A-46
                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
US-09-136-574A-2
                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                       ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
144 CGACTGCAGTACAGAGCGGCCGATACAAATGCAGCCGACAACCAGATCAAGCCGTCCTTC 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        321 GGCAGCAACATTCAGATCTCGTTTGGC----AACCATACTGGCACGAATTCGGATACG 374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6 LysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26 GlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValArgTyrTrp
                                                                                                                                                                 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 19-JAN-1991
ATFORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, Elias J. REFERENTATION NUMBER: 3425.204-US
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x US-07-862-588B-1 (1-2977)
405 Lexington Avenue, 62nd floor
                                                                                                                               COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: DNA (genomic)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.05e-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182.50
66.27%
43.37%
38.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2977 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUCLEIC ACID
DEDNESS: double
                                                                                           COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: CDS
LOCATION: 677..2776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN: NCIMB 40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |||::::::
375 TACGTGGAG 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  86 TyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OTHER INFORMATION
                                   New York
: U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                 New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS:
TOPOLOGY: lin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-07-862-588B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1, Application US/07862588B
Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Jobergensen, Per Linaa
APPLICANT: Gorgensen, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 GlyValLysValGlnTyrLysAsnAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US 08/932,571
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-5818
TELEFRA: 215-540-5818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x US-09-136-574A-1 (1-11707)
                                                                                                                             CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,574A
FILING DATE: 19-Aug-1998
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11707 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
                                                                                           COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
               COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4275 GATTATTACTTGGAG 4289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.15e-15
192.50
62.35%
43.53%
41.04%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-07-862-588B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-136-574A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
```

```
, ORGANISM: Clostridium thermocellum US-09-198-955A-11
       ; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-198-956-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lange, Niels E.
Bjornvad, Mads E.
Moller, Soren
Glad, Same O. S.
Kauppinen, Markus S.
                                                                                                                       9.37e-13
165.00
59.79%
36.08%
35.18%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 TrpAlaAlaMet-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Andersen, Lene N. Schulein, Martin
                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JS-09-198-955A-11
                                                                                                         Alignment Scores
SEQ ID NO 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: DNA
                                                                                                                                                                                                  Query Match:
                                                                                                                          Pred. No.:
                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1000 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAACTC 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1117 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels Erik K.
APPLICANT: Lange, Niels Erik K.
APPLICANT: Schnorr, Kirk
TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-US
CURRENT APPLICATION NUMBER: US/09/198,956
CURRENT APPLICATION NUMBER: 1344/97
EARLIER APPLICATION NUMBER: 60/067,240
EARLIER PILING DATE: 1997-11-24
EARLIER PILING DATE: 1997-11-24
EARLIER FILING DATE: 1997-11-26
SEARLIER FILING DATE: 1997-11-26
NUMBER OF SEQ ID NOS: 26
NUMBER OF SEQ ID NOS: 26
                                                                                                                                                                                                                                                                                                                                                                                                                    1438
35
23
29
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                        GENERAL INCE NOT STATE APPLICANT: Kauppinen, Markus APPLICANT: Kauppinen, Martin APPLICANT: Schulein, Martin APPLICANT: Schulein, Martin APPLICANT: Schulein, Martin APPLICANT: Bjornvad, Mads TITLE OF INVENTION: No. 6566114el Mannanases FILE REFERENCE: 5440.204-US CURRENT FILING DATE: 1999-06-24 NUMBER OF SEQ ID NOS: 55 SOFTWARE: Patentin version 3.1 SEQ ID NO 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-917-376-5 (1-88) x US-09-339-159B-3 (1-1438)
                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                Sequence 3, Application US/09339159B Patent No. 6566114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 9, Application US/09198956 Patent No. 6165769
                                                                                                                                                                                                                                                                                                                                                                                                                    8.98e-13
165.00
59.79%
36.08%
                                                                                                                                                                                                                                                                                                                                            ; ORGANISM: Bacillus sp. 1633
US-09-339-159B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 TrpAlaAlaMet----
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                          LENGTH: 1438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 6
US-09-198-956-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  72
                                                                                                                                                                                                                                                                                                                               TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

```
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                      991 GTATCAGGCAATTTGAAGGTTGAATTCTACAACAGCAATCCTTCAGATACTACTAACTCA 1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                           21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
     1482
35
23
29
10
Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Kaupping, Kirk APPLICANT: Schnorr, Kirk APPLICANT: Schnorr, Kirk APPLICANT: Kongsbak, Lark TITLE OF INVENTION: No. 6187580el Pectate Lyases FILE REFERENCE: 5378,200-US CURRENT APPLICATION NUMBER: US/09/198,955A CURRENT FILING DATE: 1998-11-24
                                                                                                                                                                                                US-09-917-376-5 (1-88) x US-09-198-956-9 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FastSEQ for Windows Version 4.0
                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR PELICATION NUMBER: 60/667,249
PRIOR PELICATION NUMBER: 60/667,240
PRIOR PELICATION NUMBER: 60/667,240
PRIOR PELICATION NUMBER: 09/073,684
PRIOR PELICATION NUMBER: 09/073,684
PRIOR PELICATION NUMBER: 09/184,217
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR PELICATION NUMBER: 09/184,217
PRIOR PELICATION NUMBER: 09/184,217
PRIOR PELICATION NUMBER: 09/184,217
PRIOR PELICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
SEQ ID NO: 32
SEQ ID NO: 11
LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 11, Application US/09198955A Patent No. 6187580
```

```
1168 cargérigciaraarcegcagraacegcagcracaacegaartacrrcaaarferaaagga 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40 :::::: |||||||||||| ::::::||||||||| :::
                                                                                                                                                                                                                                                                                                                                                                                  --GlyCysGlyAsnIleArgAla 71
                                                                                                                                                    1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                         41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Schulein, Martin
APPLICANT: Schulein, Martin
APPLICANT: Schulein, Martin
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels Erik K.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Schnorr, Kirk
TITLE OF INVENTION: Licheniformis
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-US
CURRENT APPLICATION NUMBER: US/09/670,141
CURRENT APPLICATION NUMBER: 09/198,956
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1999-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-26
SOFTWARE: PASLSEQ for Windows Version 3.0
SOFTWARE: PASLSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1482
35
23
29
10
     233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                 US-09-917-376-5 (1-88) x US-09-694-531-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-917-376-5 (1-88) x US-09-670-141-9 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                      61 TrpAlaAlaMet-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 9, Application US/09670141
Patent No. 6429000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: DNA ORGANISM: Bacillus licheniformis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9.37e-13
165.00
59.79%
36.08%
165.00
59.79%
36.08%
35.18%
                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-670-141-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-670-141-9
                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                           1051 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAAGTC 1110
                                                                                                                                                                                                                                     1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                          40
                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                             1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                            --- GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                      21 IlelysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                  ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :::||| |||||||||::
1228 ACMITIGIAAAATGAGTICCTCAACAAATAACGCAGACACCTACCTIGAA 1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                       1482
35
23
29
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Moller, Soren
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Schoorr, Kirk
APPLICANT: Schoorr, Kirk
APPLICANT: Schoorr, Kirk
APPLICANT: Schoorr, Kirk
APPLICANT: Schoorr, Kirk
APPLICANT: Schoorr, Kirk
APPLICANT: Sorenser, Lars
TITLE OF INVENTION: No. 6366843el Pectate Lyases
FILE REFERENCE: 5378_200_US
CURRENT APPLICATION NUMBER: US/09/694,531
CURRENT FILING DATE: 1997-11_24
PRIOR PILING DATE: 1997-11_24
PRIOR FILING DATE: 1997-11_24
PRIOR FILING DATE: 1997-11_24
PRIOR FILING DATE: 1997-11_24
PRIOR PPLICATION NUMBER: 60/067,249
PRIOR FILING DATE: 1997-12_02
PRIOR FILING DATE: 1997-12_02
PRIOR FILING DATE: 1997-12_02
PRIOR FILING DATE: 1997-12_02
PRIOR PRILING DATE: 1997-12_02
PRIOR PRILING DATE: 1997-12_02
PRIOR PRILING DATE: 1998-11_02
PRIOR FILING DATE: 1998-11_02
                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                        US-09-917-376-5 (1-88) x US-09-198-955A-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
TYPE: DNA
ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 11, Application US/09694531 Patent No. 6368843 GENERAL INFORMATION: APPLICANT: Andersen, Lene N.
                                                                                                                                                                                                                                                                                                                                                                                                                                            61 TrpAlaAlaMet-----
                                           9.37e-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Schulein, Martin
Lange, Niels E.
Bjornvad, Mads E.
                                                        165.00
59.79%
36.08%
35.18%
                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
Pred. No.:
                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-694-531-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-694-531-11
                                                                                                                                                                                                                                                                                                                                                                    41
                                           Pred. No.:
                                                            Score:
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                           원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
```

```
|||||||:::::
|168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor GITY: New York STATE: New York COUNTRY: U.S.A.
                                                                                                       72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              An Enzyme Exhibiting Cellulase Activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWAREN PARTICIPATION OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728 REGISTRATION NUMBER: 30,335 / 33,728 TELEPRENCE/DOCKET NUMBER: 3425.204-US TELEPHONE: 212 867 0123 TELEPHONE: 212 867 0123 TELEPHONE: 212 867 0298 INFORMATION FOR SEQ ID NO: 6: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 6, Application US/07862588B
Patent No. 5916796
GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Hansen, Christian
TITLE OF INVENTION: An Enzyme Exhib
WUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 59167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .17e-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 1624 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein FRAGMENT TYPE: N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120.50
54.12%
32.94%
25.69%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN: NCIMB 40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCATION: 30..1607
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                        US-07-862-588B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JS-07-862-588B-6
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1051 ATCAATCCTCAGTTCAAGGTTACTAATACCGGAAGCAGTGCAATTGATTTGTCCAAACTC 1110
                                                                                                                                                                                                                                             |||||||:::::::
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                       1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                       --GlyCysGlyAsnIleArgAla 71
                                                                    ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                               72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Algad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Solo-105
CURRENT APPLICATION NUMBER: 105/09/198,955
PRIOR FILING DATE: 1998-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR PILING DATE: 1997-12-02
PRIOR PILING DATE: 1999-12-02
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-01-02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x US-10-072-152-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 11, Application US/10072152 Patent No. 6677147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.37e-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Andersen, Lene N. APPLICANT: Schulein, Martin APPLICANT: Lange, Niels E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165.00
59.79%
36.08%
35.18%
                                                                                                                                                                                                    61 TrpAlaAlaMet---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-10-072-152-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: DNA
                                                                    41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                 8
                                                                                                                               셤
```

```
FELECOMMUNICATION INFORMATION
NCIMB 40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10036
                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                        NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-048-164A-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CITY: N
STATE:
                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                   ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc. STREET: 405 Lexington Avenue, 62nd floor
                                                                                                                                                                                                                                        61
                                                                                                                                                         LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                      62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                           2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                        42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
APPLICANT: Joergensen, Per Linaa
APPLICANT: Sch lein, Martin
APPLICANT: Sch lein, Martin
APPLICANT: Answerian
TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Zelson, Steve T. / Lambiris, Elias J. REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 3425.204-US
TELECOMMUNICATION INFORMATION: TELEPHONE: 212,867,0123
TELEPHONE: 212,867,0298
                                      US-09-917-376-5 (1-88) x US-07-862-588B-6 (1-1624)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
APLICATION NUMBER: PCT/DK91/00013
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATURE SYSTEM: PC-DOS/MS-DOS
COFFMANDED: COMPUTED STATEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 5, Application US/07862588B Patent No. 5916796
                                                                                                                                                                                                                                                                                                                                                                                                                                        1554 AATACGGATACTTAC 1568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Bacillus lautus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                   82 ThralaAspThrTyr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 1775 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New York : U.S.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY: New York
STATE: New Yor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-07-862-588B-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                원
                                                                                                                                                           ò
                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                        ò
```

```
:::
1443 CTINNNNNNNNNNNAAAAGAC---GGACCTGCGGATATGAGCTGCTCGATCGGCTGG 1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1500 GCGCAAATCGGCCGAACGAATGTTCTGCTGGCATTC----GCTAACTTTACCGGGAGT 1553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                                                                     2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 LysproglyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                 aa:
                                                                                                                                                                                1775
28
18
36
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,164A
FILING DATE: 14-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leelie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-003
                                     /evidence= EXPERIMENTAL
/transl_except= (pos: 1446
OTR)
                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
ITILE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
ADDRESSEE: PENNIE & EDMONDS
                                                                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x US-07-862-588B-5 (1-1775)
                                                                                                                                                                                                       Matches:
              OD: experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Application US/08048164A Patent No. 5496934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1554 AATACGGATACTTAC 1568
                                                                                                                                                                                1.32e-06
120.50
54.12%
32.94%
25.69%
LOCATION: 30..(1625.1775)
IDENTIFICATION METHOD: ex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 ThrAlaAspThrTyr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                     CTHER INFORMATION:
CTHER INFORMATION:
CTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New York
: New York
RY: U.S.A.
```

```
Misrock, S. Leslie
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: concur CLASSIFICATION: 435
                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: New York
STATE: New York
                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  U.S.A.
                                                                                                                                                                                                            TOPOLOGY: unl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: U
ZIP: 10036
                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                  US-08-048-164A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-460-462-1
                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   184 GCTGGTGCATTATTAGGAATTAGCTATGTTGATAACACTAGCAAAGTGACAGCAAACTTC 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 AlaAlaMetGlyCysGlyAsn------11eArgAlaSerPhe 73
                                                                                                                                                                                                                                                                                                                                                                                                                                          2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                  486
22
34
12
44
12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Shoseyov, Oded
APPLICANT: Store Sources: 21
CORRESSED NOBNOES ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x US-08-048-164A-1 (1-486)
                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: US/08/048,164A
FILING DATE: 14-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 3, Application US/08048164A Patent No. 5496934
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
                                                  INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                 114.00
52.58%
29.90%
24.31%
                                                                                                                                           unknown
                                                                                                                                       TOPOLOGY: unknown MOLECULE TYPE: DNA
                                                                                                                                                                                                               1..486
                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
US-08-048-164A-3/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10036
                                                                                                                                                                                             NAME/KEY:
                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                               LOCATION:
                                                                                                                                                                                                                              US-08-048-164A-1
                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
```

```
303 GCTGGTGCATTATTAGGAAATAGCTATGTTGATAACACTAGCAAAGTGACAGCAAACTTC 244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               420 ACACCAATAATCAAATTACTAACACATCTGACAGTGATTTAAATTTAAATGACGTAAAA 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/08460462
Patent No. 5670623
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shoised, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspleuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----IleArgAlaSerPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               243 GTTAAAGAAACAGCCAA---ACATCAACCTATGATACATGTTGAA 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                           486
22
34
12
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM PC Comparible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x US-08-048-164A-3 (1-486)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 AlaAlaMetGlyCysGlyAsn-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUTING DATE: CONCERN APPLICATION NUMBER: US/08/460,462
APPLICATION NUMBER: US/08/460,462
FILING DATE: CONCURRENTLY herewith
                            7809-003
                                                            TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
REGISTRATION NUMBER: 18,872
REPERENCE/DOCKET UNBER: 78
TELECOMMUNICATION
TELEPHONE: (212) 790-9090
                                                                                                                                                                                                                                                                                                                                                                              64e-06
                                                                                                                                                                                                                                                                                                                                                                                                 114.00
52.58%
29.90%
24.31%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
```

```
67 ACACCAATAATCAAAATTACTAACACATCTGACAGTGATTTAAATTTAAATGACGTAAAA 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22 LysproGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----IleArgAlaSerPhe 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               244 GTTAAAGAAACAGCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                        486
222
24
124
124
                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-5 (1-88) x US-08-460-462-1 (1-486)
                           PILING DATE: 14-APR-1933
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9990
TELEFA: (6141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 base pairs
TYPE: mucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 AlaAlaMetGlyCysGlyAsn-----
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
                                                                                                                                                                                                                                                                                                                                                                                          1.64e-06
                                                                                                                                                                                                                                                                                                                                                                                                       114.00
52.58%
29.90%
24.31%
                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                                                                         TOPOLOGY: unknown MOLECULE TYPE: DNA FEATURE:
                                                                                                                                                                                                                                                                                                                          1..486
                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                     ; NAME/KEY:
; LOCATION:
US-08-460-462-1
                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
```

Search completed: May 12, 2004, 15:15:30 Job time : 26.976 secs

Appli Appli Appli Appli Appli

Title: Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Total number

Searched:

```
Sequence 684, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 1529, A
Sequence 5, Appli
Sequence 17, Appl
Sequence 1716, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 742, App
Sequence 222, App
Sequence 11, Appli
Sequence 21, Appli
Sequence 21, Appl
Sequence 188, Appl
Sequence 188, Appl
Sequence 19, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 85, Appl
Sequence 10, Arri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 393, App
Sequence 393, App
Sequence 18, Appl
Sequence 18, Appl
Sequence 51043, A
Sequence 1, Appli
Sequence 85, Appli
                                                                                                                                                                          Sequence 46838, A
Sequence 3, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 21, App
Sequence 9462, A
Sequence 3, Appli
                                                                                                                                                                                                                        Sequence 11,
Sequence 11,
                                                       Sequence 2,
Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: ADNEXEN, TODD B.
APPLICANT: DECKER, STEPHEN R.
APPLICANT: HIMMEL B.
TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFREENCE: 40197.70501
CURRENT APPLICATION NUMBER: US/09/917,378
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 2289
                                                                                                                                                      Sequence
                                                                                                      Sequence Sequence
10 US-09-917-378-2
110 US-09-917-378-2
113 US-09-917-384-2
115 US-10-155-400-2
115 US-10-156-401-2
116 US-10-156-401-550
119 US-10-156-401-550
119 US-10-156-401-31
119 US-10-156-401-31
119 US-10-156-401-31
119 US-10-166-401-31
119 US-10-166-401-31
119 US-10-166-401-31
119 US-10-166-401-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-166-51-31
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-116-802-38
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
119 US-10-117-722-393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-074-095-85
US-10-074-095-85
US-10-212-872-85
US-10-617-334-21
US-10-617-334-21
US-10-452-510-21
US-07-822-514-9462
US-09-822-268A-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15 US-10-156-761-3
9 US-09-764-860-85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/09917378 Publication No. US20030119093A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                         11146
95147
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
96187
                                                                                                                                                                                                                                                1482
1314
4059
4056
4452
768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: DING, SHI-YOU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16.0
15.6
15.5
15.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-378-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                             155.5
150.5
130.5
                                                                                                                                                                                                                                                                                                                                                                                                                 78.5
                                                                                                                                                                                                                          Command line parameters:
-MODEL=frame+p2.model=1.DEV=xlh
-MODEL=frame+p2.model=1.DEV=xlh
-De-Cgnz 1/USPTO spool/USO9917376/runat 11052004_114537_28336/app_query.fasta_1.1429
-De-Cgnz 1/USPTO spool/USO9917376/runat 11052004_114537_28336/app_query.fasta_1.1429
-De-De-Dublished Applications NA -QFNT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -RND=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXIEN=200000000 -USER=USO9917376_@CGN 1 1 809_@runat 11052004_114537_28336
-NCPU=5 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -MIT -DSPBLOCK=100
-LONGLOG -DBV TIMEOUT=120 -WARRD TIMEOUT=30 -THRRANS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
                                                                                                                                                    ; Search time 135.119 Seconds (without alignments) 2950.286 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                            1 VSGGVKVQYKNNDSAPGDNQ.....IRASFGSVNPATPTADTYLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                           - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2941586 seqs, 2264995651 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Published Applications NA:*
                                                                                                                                                                                                                                                                                                                                                              Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                    May 12, 2004, 09:26:52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query
Match Length DB
                                                                                                                                                                                                                                                US-09-917-376-5
                                                                                                                                                                                                                                                                                                                                             BLOSUM62
```

Jatabase

Score

Result Š.

₽ g Dp ò D ò g à 유

√ò

```
1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGTCGTCGACACTGGTGTACAACTGTGAC 1929
                                                                                                                                           1930 TGGGCGGCGATGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAGAACAATGATTCGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1810 ATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGGTGGATTTGTCGACGGTG 1869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1870 ACGGTGCGGTACTGGTTCACCCGGGATGGTCGTCGTCGACACTGGTGTACAACTGTGAC 1929
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1930 TGGGCGCGATGGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGAACCCGGCGAGG 1989
                   ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                   TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                      Sequence 2, Application US/09917383
; Sequence 2, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: DECKEN, WILLIAM S.
; APPLICANT: HIWMEL, MICHAEL E.
; APPLICANT: HIWMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION: CELLULOLYTICUS
; TITLE OF INVENTION UNDER: US/09/917,383
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 3.687
; MUMCH: 3.687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: Description of Artificial Sequence: Segment of OTHER INFORMATION: GuxA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3687
87
1
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x US-09-917-383-2 (1-3687)
                                                                                                                                                                                                          1990 CCGACGCGGACACCTGCAG 2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88
                                                                                                                                                                                     81 ProThrAlaAspThrTyrLeuGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
US-09-917-376-2
Sequence 2, Application US/09917376
; Publication No. US20040038334A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: DNA ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.13e-55
466.00
100.00%
98.86%
99.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-917-383-2
                       41
                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                     8
                                                                                                                                                                                                    1483 ACGGGGGAACTGGTTCACCCGGGATGGTGGGTCGTCGACACTGGTGTACAACTGTGAC 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1543 TGGGCGCGCGATGGGGGTGTGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 1602
                                                                                                                                                                                                                                                                                                         ATCAAACCGGGTCTCCAGTTGGTGAATACGGGGTCGTCGTCGTGGATTTGTCGACGGT 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1750 GTGTCGGGTGGGCTGAAGGTGCAGTACAAGAACAATTGGGCGCCGGGTGATAACCAG 1809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1810 ATCAPACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGTCGATTTGTCGACGGTG 1869
                                                                                                                                                                                                                                                                                  21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal 40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                   TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 IlelysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/09917384

| Sequence 2, Application No. US20030096342A1
| Publication No. US20030096342A1
| GENERAL INFORMATION:
| APPLICANT: DING, SHI-YOU
| APPLICANT: ADNEY, WILLIAM S.
| APPLICANT: VINZANT, TODD B.
| APPLICANT: HIMMEL, MICHAEL B.
| APPLICANT: HIMMEL, MICHAEL B.
| TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS FILE REFERENCE: 40170.6US01
| CURRENT APPLICATION NUMBER: US/09/917,384
| CURRENT FILING DATE: 2001-07-28
| NUMBER OF SEQ ID NOS: 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURE:

OTHER INFORMATION: GuxA
US-09-917-384-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3687
87
1
0
0
              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-917-376-5 (1-88) x US-09-917-384-2 (1-3687)
                                                                                                                                                              US-09-917-376-5 (1-88) x US-09-917-378-2 (1-2289)
                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1603 ccaacacacacacaracaracaa 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: DNA ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.13e-55
466.00
100.00%
98.86%
99.36%
                2.21e-56
469.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                         Percent Similarity:
Best Local Similarity:
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO 2
LENGTH: 3687
                                                                                                                                                                                                                                                                                                                          1423
                                                                                                                                                                                                                                                                                                                                                                   41
                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
```

40

80

셤 8

ò

```
.. (2223)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , NAME/KEY: CDS
, LOCATION: (1)
US-10-156-761-550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -10-156-761-550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
Pred. No.:
                                                                                                       Alignment Scores:
                                                                                                                                                                                                             Query Match:
                                                                                                                                Pred. No.:
                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2665 ATCAAGCCGGGTTTGCAGGTGGTGAATACCGGGTCGTCGTCGTCGATTTTGTCGACGGTG 2724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2725 ACGGIGCGGTACTGGTTCACCCCGGGATGGTGGTCTCGTCGACACTGGTGTACAACTGTGAC 2784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2785 TGGGCGGCGATCGGGTGTGGGGAATATCCGCGCCTCGTTCGGCTCGGTGAACCCGGCGACG 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/10155400

Sequence 2, Application US/10155400

Publication No. US20030108988A1

GENERAL INFORMATION:

APPLICANT: DING, SHI-YOU

APPLICANT: UNZANT, TODD B.

APPLICANT: HIMMEL, MICHAEL B.

TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS

TITLE OF INVENTION: CELLULOLYTICUS

FILE REPERENCE: NREL 01-36A

CURRENT APPLICATION NUMBER: US/10/155,400

CURRENT FILING DATE: 2002-10-22

NUMBER OF SEQ ID NOS: 7
GENERAL INFORMATION:
APPLICANT: DING, SH-YOU
APPLICANT: ADNEY, WILLIAM S.
APPLICANT: VINANT, TODD B.
APPLICANT: VINANT, TODD B.
APPLICANT: HIMMEL, MICHAEL E.
TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
TITLE OF INVENTION: CELLULOLYTICUS
FILE REFERENCE: 40197.41501
CURRENT APPLICATION NUMBER: US/09/917,376
CURRENT FILING DATE: 2001-07-28
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATCHLIN VET. 2.1
SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                     other or unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-917-376-5 (1-88) x US-09-917-376-2 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2845 CCGACGCGGACACCTACCTGCAG 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: DNA
ORGANISM: Acidothermus cellulolyticus
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                               TYPE: DNA ORGANISM: Acidothermus cellulolyticus
                                                                                                                                                                                                                                                                                                                                                                                                       ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.98e-55
462.00
100.00%
97.73%
98.51%
                                                                                                                                                                                                                                                                                                                                                                             ; LOCATION: (2869)
; OTHER INFORMATION: a, c, t,
US-09-917-376-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: modified_base
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                        LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH: 2869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-10-155-400-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                       FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
2785 TGGGCGCGATCGGGTGTGGGAATATCCGCGCCTCGTTCGCTCGGTGAACCCGGCGACG 2844
                                                                                                                                                                                                                                                                                                                                                                                                                                               2724
                                                                                                                                                                                                                                                                                                                                              2605 GTGTCGGGTGGATGATGCAGTATAAAGAATAATGATTCGGCGGGTGATAATCAG 2664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2725 ACGGTGCGGTACTGGTTCACCCGGGATGGTGGTCGTCGTCGACACTGTACAACTGTGAC 2784
                                                                                                                                                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                    1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                             2665 ATCAAGCCGGGTTTGCAGGTGGAATACCGGGTCGTCGTCGTCGTCGTGGATTTGTCGACGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                       21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2223
49
13
25
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-10-155-400-2
                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x US-10-155-400-2 (1-2869)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: INEDA, HARUO
APPLICANT: INEDA, HARUO
APPLICANT: INEDA, HARUO
APPLICANT: INEDA, HARUO
APPLICANT: SHIRAMA, JUN
APPLICANT: SHIRAM, HROSHI
APPLICANT: SHIRA, TADAYOSHI
APPLICANT: SHIRA, TADAYOSHI
APPLICANT: HATTORI MASAHIRA
ITILE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-226
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: UP 2001-204089
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-05-30
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 550
LENGTH: 2223
                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2845 ccgaccccacaccraccraccac 2868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81 ProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 550, Application US/10156761 Publication No. US20030119018A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: DNA ORGANISM: Streptomyces avermitilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3.69e-26
257.00
71.26%
56.32%
54.80%
                                                                                                                          2.98e-55
462.00
100.00%
97.73%
98.51%
```

```
APPLICANT: Cao, Yongwei
APPLICANT: Cao, Yongwei
APPLICANT: Black, Gregory J.
APPLICANT: Alack, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10 (5205.2)
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR FILLING DATE: 2002-02-21
                                                        706336 GCCGGCTACTACTTCACCCGGGACAGCGGCTCGCCCACCGTGAACGCCTGGTGCGACTAC 706395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :::||||||||||||:::
1255 CAGATTGGATGCGGCAATGTGACACACAGTTTGTGACGTTGCATAAACCAAAGCAAGGT 1314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       63 AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             44 TyrTrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAla
                                       AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-917-376-5 (1-88) x US-10-369-493-46838 (1-1527)
                                                                                                                                                    706456 GGAGCCGACCTACCTCGAA 706476
                                                                                                                                                                                                                               Sequence 46838, Application US/10369493
Publication No. US20030233675A1
                                                                                                                 82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Sequence 3, Application US/10372054; Publication No. US20030203466A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1315 GCAGATACCTATCTGGAA 1332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83 AlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .02e-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Kauppinen, Markus
APPLICANT: Schulein, Martin
APPLICANT: Schnorr, Kirk
APPLICANT: Andersen, Lene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: DNA ORGANISM: Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     190.00
65.12%
40.70%
40.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 46838
LENGTH: 1527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                           US-10-369-493-46838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JS-10-369-493-46838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-372-054-3
                                                                          셤
                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||||||::||||||||::
| GCCGCTACTACTACTTCACCCGGGACAGCGGCTCGCCCACCGTGAACGCCTGGTGCGACTAC 1950
                                                                                                               41
                                                                                                                                                                                                                                                                62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                     SerGlyGlyValLygValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                         ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                               22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      449
113
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-5 (1-88) x US-10-156-761-1 (1-9025608)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          other or unknown
US-09-917-376-5 (1-88) x US-10-156-761-550 (1-2223)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
CURENT APPLICATION NOWER: US/10/156,761
CURENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR PILING DATE: 2001-05-30
PRIOR PILING DATE: 2001-06-02
NUMBER OF SEQ ID NOS: 15109
                                                                                                                                                                                                                                                                                                                                                                    TYPE: DNA ORGANISM: Streptomyces avermitilis
                                                                                                                                                                                                                                                                                                                                          82 ThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1, Application US/10156761 Publication No. US20030119018A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ຫ້
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.72e-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  257.00
71.26%
56.32%
54.80%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; OTHER INFORMATION: a, t, US-10-156-761-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: OMURA, SATOSHI APPLICANT: IKEDA, HARUO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: misc feature LOCATION: (4187715)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LENGTH: 9025608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    엄
                                                                      g
                                                                                                                                                셤
                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                     吕
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                               8
                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                          ò
```

```
Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-10-072-152-11
                                                                                                                                                                     US-10-655-433-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               심
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 940 GTATCAGGCAATTTGAAGGTTGAATTCTACAACAGCAATCCTTCAGATACTACTAACTCA 999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60
                                                                                                                                                                                                                                                                                                                                                                                                                             yalserglyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Andersen, Lene N.
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Bjornvad, Mads E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Rauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, No. US20040067572Alel Pectate Lyases
TITLE OF INVENTION: No. US20040067572Alel
FILE REFERENCES: 25792.200-US
CURRENT FILING DATE: 1998-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-11-24
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
                                                                                                                                                                                                                                                                       1438
35
23
29
10
                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x US-10-372-054-3 (1-1438)
            TITLE OF INVENTION: Novel Mannanases
FILE REFERENCE: 5440 204-US
CURRENT APPLICATION NUMBER: US/10/372,054
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US/09/339,159B
PRIOR PILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 55
SEQ ID NO 3
SEQ ID NO 3.1
                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 11, Application US/10655433
Publication No. US20040067572A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TrpAlaAlaMet-------
                                                                                                                                                                                                                                                                         2.72e-13
                                                                                                                                                                                                   ORGANISM: Bacillus sp. 1633
                                                                                                                                                                                                                                                                                        165.00
59.79%
36.08%
35.18%
Bjornvad, Mads
                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                   LENGTH: 1438
                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
US-10-655-433-11
                                                                                                                                                                                                                      US-10-372-054-3
                                                                                                                                                                                                                                                                                                                                                                                                                             Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41
APPLICANT:
                                                                                                                                                                                       TYPE: DNA
                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
||||||||:::
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41 ThrvalArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --- GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21 ileLysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :::||| |||||||||:::
1228 ACMITIGTAAAATGAGTICCICAACAAATAACGCAGACACCTACCTIGAA 1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Glad, Sanne O. S.
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Kirk
APPLICANT: Markus S.
APPLICANT: Markus S.
APPLICANT: Markus S.
APPLICANT: Markus S.
APPLICANT: Markus S.
APPLICANT: Markus S.
APPLICANT: Markus S.
TITLE OF INVENTION: No. US20020142438Alel Pectate Lyases
FILE REPERENCE: 5378.200.405
CURRENT APPLICATION NUMBER: US/10/072,152
CURRENT FILING DATE: 2002-02-07
                                                                                                                                                                                                                                                                                                                                       1482
35
23
29
10
                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-917-376-5 (1-88) x US-10-655-433-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARLE FILING DATE: 2002-02-07

PRIOR APPLICATION NUMBER: US/09/198,955

PRIOR FILING DATE: 1998-11-24

PRIOR FILING DATE: 1997-11-24

PRIOR FILING DATE: 1997-11-24

PRIOR PRING DATE: 1997-11-24

PRIOR PRING DATE: 1997-11-24

PRIOR PRING DATE: 1997-12-02

PRIOR FILING DATE: 1997-12-02

PRIOR FILING DATE: 1997-12-02

PRIOR FILING DATE: 1997-12-02

PRIOR APPLICATION NUMBER: 60/067,240

PRIOR PRING DATE: 1997-12-02

PRIOR APPLICATION NUMBER: 09/073,684

PRIOR FILING DATE: 1998-05-06

PRIOR PILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1988-05-06
PRIOR PILING DATE: 1988-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: RastSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 11, Application US/10072152
Publication No. US20020142438A1
GENDEAD INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
                                                                                                                                                                                                               TYPE: DNA ORGANISM: Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                       2.84e-13
165.00
59.79%
36.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 TrpAlaAlaMet-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Lange, Niels E.
Bjornvad, Mads E.
Moller, Soren
```

```
5 VallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGly 24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                US-09-784-554B-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JS-09-784-554B-1
                                                          25
                                                                                      938
                                                                                                                 45
                                                                                                                                                                      64
                                                                                                                                                                                                                             84
                                                                                                                                                                                                                                                       1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                          à
                                                                               셤
                                                                                                                ઠે
                                                                                                                                          셤
                                                                                                                                                                      ò
                                                                                                                                                                                             g
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                      |||||||:::
1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAATGTAAAAGGA 1227
                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                     ---GlyCysGlyAsnIleArgAla 71
                                                                                                                                                                                                                                                                    1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
                                                                                                                                                                                                                                                                                                                                                                              41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 684, Application US/09974300
; Battent No. US20020146721A1
; GENERAL INFORMATION:
    APPLICANT: BERKA, Randy M.
    APPLICANT: Clausen, ID Groth
    TITLE OF INVENTION: Methods For Monitoring Multiple Gene
    TITLE OF INVENTION: Expression
    FILE REFERENCE: 10085.500-US
    CURRENT APPLICATION NUMBER: US/09/974,300
    CURRENT FILING DATE: 2001-10-05
    PRIOR APPLICATION NUMBER: 09/680,598
    PRIOR FILING DATE: 2000-10-06
    PRIOR FILING DATE: 2000-10-06
    PRIOR FILING DATE: 2000-10-05
    NUMBER OF SEQ ID NOS: 8481
    SOFTWARE: FastSEQ for Windows Version 4.0
    SEQ ID NOS: 8481
    SEQ ID NOS: 8481
                                                                                                                                        1482
35
23
29
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1314
32
16
33
4
                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-917-376-5 (1-88) x US-09-974-300-684 (1-1314)
                                                                                                                                                                                                                                       US-09-917-376-5 (1-88) x US-10-072-152-11 (1-1482)
                                                                                                                                                                                                                                                                                                                                                                                                                                      61 TrpAlaAlaMet------
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1482
                                                                   TYPE: DNA; Clostridium thermocellum US-10-072-152-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: DNA ORGANISM: Bacillus licheniformis
                                                                                                                                        2.84e-13
165.00
59.79%
36.08%
35.18%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3".34e-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               157.00
56.47%
37.65%
33.48%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
Query Match:
                                                                                                                                                                                   Best Local Similarity:
                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-974-300-684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-974-300-684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 1314
                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . No. :
                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                            엄
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                   g
G
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
998 rigitataaakigaatigicaaaggacaaat------titgacrigigakitargicicaa 1048
                                                                                                                                                                                                                                                                            |||::: :::|||||||:: ||||||3598 GACTGCTGCTGCTCAATACCG 3657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3718 TACTACTACACAATCGACGGGACCGTGAGCAG---ACTTTCAACTGCGACTATGCGACG 3774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ::: |||
1775 CTGAGCTGCTCAAAGCTGAAAGCTGGTTAAAATGGAGAAGGCTGCAACCGGTGCC 3834
::: ||||||||||:::|||
878 ATAGCAGTACAATACAGAGCGGGGGACAATGTAAACGGCAACCAATCCGCCCTCAG 937
                                                                                          TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                         LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 GlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24 GlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4059
27
21
36
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Schootr, Kirk
APPLICANT: Schootr, Fer Lina
APPLICANT: Scholin, Martin
APPLICANT: Scholin, Martin
TITLE OF INVENTION: PAMILY 44 XYLOGLUCANASES
FILE REFERENCE: 10017.200-US
CURRENT FAPLICATION NUMBER: US/09/784,554B
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
LENGTH: 4059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-917-376-5 (1-88) x US-09-784-554B-1 (1-4059)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1, Application US/09784554B Publication No. US20030032162A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 14
US-09-784-554B-3
; Sequence 3, Application US/09784554B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                       GACACGTATCTTGAA 1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ||| ||||||:::
3835 GATTATTTTGGAA 3849
                                                                                                                                                                                                                                                                                                                                            AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2.5e-11
155.50
56.47%
31.76%
33.16%
```

```
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3715 TACTACTACACGATCGACGGTGACCGTGAGCAG---ACATTCAACTGCGGATGCGGTG 3771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3772 CTGAGCTGCTCGAAGCTGAATGGTAAGCTGGTTAAAATGGATAAAGCTGCAACCGGTGCT 3831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 GlyValLysValGhnTyrLysAsnAsnAspSerAlaProGlyAspAsnGhnIleLysPro 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: Identification of Essential Genes in Microorganisms FILE REFERENCE: ELITRA.034A
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR APPLICATION NUMBER: 60/207,727
                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-917-376-5 (1-88) x US-09-784-554B-3 (1-4056)
              GENERAL INFORMATION:
APPLICANT: Schnorr, Kirk
APPLICANT: Schnorr, Fartin
TITLE OF INVENTION: FALTIN
TITLE OF INVENTION: FALTIN
FILE REFERENCE: 10017, 00-018
CURRENT APPLICATION NUMBER: US/09/784,554B
CURRENT FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 16
SEQ ID NO 3
SEQ ID NO 3
                                                                                    LY 44 XYLOGLUCANASES
'00-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 15299, Application US/10282122A Publication No. US20040029129A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                          ORGANISM: Paenibacillus polymyxa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3832 GATTATTATTIGGAA 3846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Malone, Cheryl
APPLICANT: Olisen, Kari
APPLICANT: Olisen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Carr, Grant
APPLICANT: Fawick, John
APPLICANT: Fawick, Sobert
APPLICANT: Forsyth, R.
                                                                                                                                                                                                                                                                                                                                              1.29e-10
150.50
56.47%
31.76%
32.09%
Publication No. US20030032162A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84 AspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-10-282-122A-15299
                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                         LENGTH: 4056
                                                                                                                                                                                                                                                                                                US-09-784-554B-3
                                                                                                                                                                                                                                                              TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               원
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
              PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-09-06
PRIOR PAPLICATION NUMBER: 60/230,347
PRIOR PILING DATE: 2000-09-09
PRIOR PILING DATE: 2000-09-09
PRIOR PILING DATE: 2000-02-23
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR PELLING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR PILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR PILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
PRIOR FILING DATE: 2001-02-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 SerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 LysProGlyLeuGlnLeuValAsnThrGlySerSerSerValAspLeuSerThrValThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 AlaAlaMetGlyCysGly------AsnIleArgAlaSerPhe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74 GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4452
30
23
33
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-917-376-5 (1-88) x US-10-282-122A-15299 (1-4452)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Clostridium acetobutylicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: May 12, 2004, 18:28:26
Job time : 827.119 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.04e-07
130.50
55.79%
31.58%
27.83%
2000-02-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-282-122A-15299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
```

DATE 5/14/07	APPLICATION NUMBER 09/917,376
DOC CODE	DOC DATE 10/6/05

DELIVER THE ATTACHED FIFLE/DOCUMENT TO THE TC SCANNING CENTER

CONTRACTOR: THE ATTACHED FILE/DOCUMENT MUST BE INDEXED AND SCANNED INTO IFW WITHIN 8 WORK HOURS; UPLOADING OF THE SCANNED IMAGES SHOULD OCCUR NO LATER THAN 16 WORK HOURS FOLLOWING RECEIPT OF THIS REQUEST

AFTER SCANNING, ORIGINAL DOCUMENTS SHOULD BE BOXED IN ACCORDANCE WITH INSTRUCTIONS